## Title
Non-native contrasts in Tongan loans



## Journal
Phonology, 36(1)

## ISSN
0952-6757

## Authors
Zuraw, K
O'Flynn, KC
Ward, K

## Publication Date
2019-02-01

## DOI
10.1017/S095267571900006X

Peer reviewed

## Non-native contrasts in Tongan loans*

## 1 Introduction

A central issue in loanword phonology is how loan adapters perceive non-native contrasts, and how this affects the form of the borrowed word.

On one side, researchers such as Dupoux and colleagues emphasise the difficulty that even advanced learners have perceiving non-native contrasts. For example, Dupoux et al. (1997) and Dupoux, Peperkamp and Sebastián-Gallés (2001) document 'stress deafness' in French speakers: French lacks contrastive stress, and French speakers perform poorly at recognising or representing stress. This poor performance holds for those who have learned Spanish (which has contrastive stress) to an advanced level (Dupoux et al. 2008)—and even for many simultaneous Spanish-French bilinguals (Dupoux, Peperkamp & Sebastián-Gallés 2010). Work on vowel/zero contrasts has shown similar results. In Japanese, most consonants must be followed by a vowel; therefore, there is no contrast between CC and CVC, for most choices of CC. Accordingly, Japanese speakers tend to perceive …CC… stimuli (e.g., [ebzo]) as …CVC… ([ebuzo]) (Dupoux, Kakehi, et al. 1999, Dupoux, Fushimi, et al. 1999). Peperkamp & Dupoux (2003) and Peperkamp (2005) argue that these 'illusory vowels' explain epenthesis in loans from English. Rather than starting with an English-like form /sfɪŋks/ 'sphinx' and inserting vowels to conform to Japanese phonology ([s**u**fiŋk**u**s**u**]), a Japanese speaker perceives the English word as something like /sufiŋkusu/ from the beginning.

On the other side, authors such as LaCharité and Paradis (2002) have emphasised the degree to which bilinguals *can* perceive non-native contrasts. After all, subjects in the studies cited above do perform above chance under certain conditions, and the most-influential loan adapters might be the most-proficient bilinguals, with the best ability to discriminate second-language contrasts. Under this view, although Japanese adapters of English words might make some perceptual errors, on the whole they will still tend to perceive cluster-initial /pleɪ/ 'play' and /skaɪ/ 'sky' differently from CV-initial /pʊɾɪŋ/ 'pudding' and /sədæn/ 'sedan'. In this example, from Kubozono (2006: 1147), all four words are adapted into Japanese as /CVC…/, but their accentuation patterns differ. Light-heavy loans tend to have initial accent—/púriN/, /sédaN/—unless the initial vowel is epenthetic—/purée/, /sukái/. Even if initial adapters perceive /pl/ and /sk/ as having some sort of weak vowel, or if they perceive them as CC but don't quite produce them as such, some difference between /CC/ and /CVC/ could still be preserved. Davidson (2006, 2010) shows that when English speakers are asked to imitate a CC sequence that is illegal for them, they often produce a 'transitional vowel', yielding a sequence transcribable as [Cᵊ C], acoustically different from an underlying /CəC/.

* We thank Manu Tu'uholoaki, Saia Moala, and Piula Tonga for their expertise in the Tongan language and their patience in assisting with this project. Other participants in the UCLA linguistics field methods course also helped shape this project: Hilda Koopman, Natasha Abner, Nicki Acker, Byron Ahn, Jason Bishop, Heather Burnett, Marc Garellek, Grace Kuo, James Pannacciulli, Craig Sailor, Nancy Finifter, and James White. Audiences at UCLA, UC Berkeley, UMass, Stanford, and Seoul National University provided valuable feedback, especially John Kingston and Jongho Jun. We are grateful to the UCLA statistics consulting staff for their help, especially Andy Lin, Phil Ender, Christine Wells, and Xiao Chen. Funding from the UCLA Academic Senate's Committee on Research supported this research.

We present three case studies of contrast in Tongan loans from English: secondary stress, vowel deletion, and final vowel length. Our first point is that loans display contrasts not present in native Tongan words. As in French, stress is predictable in native Tongan words; but secondary stress position is unpredictable in long English loans ([**mò**nokaláme] 'monogram' vs. [ko**lò**nitíni] 'quarantine') (Schütz 2001). Like Japanese, Tongan has no contrast between CC and CVC (and without the exemptions for nasals and geminates that Japanese allows). Also similarly to Japanese, Tongan vowels sometimes devoice or delete, encouraging CC and CVC to be perceived as belonging to the same category. But, we show that certain vowels are more susceptible to deletion than others ([kàs_tomáa] 'customer', where all three of our consultants produced deletion, vs. [tipòs**i**tóa] 'depositor', where none did). We analyze both secondary stress and vowel deletion as reflecting a contrast between full and weak vowels. We deem the contrast marginal because it is realised either indirectly (through secondary stress, or length of other vowels) or in rates of variation rather than clear minimal pairs (vowel deletion).[1]

Our second point is that loans reflect sensitivity to English contrasts. Secondary stress, vowel deletion, and final vowel length all treat English CC differently from English CVC. And while final vowel length is contrastive in Tongan, we show that in loans it depends on English stress.

After we present these three cases (sections 3, 4, 5), we offer an analysis in the multilevel model of Boersma (1998) and Boersma and Hamann (2009). Following Broselow (2009), we argue that the Tongan loan patterns can be captured using only constraints that are plausibly active for native words too, including constraints that reflect perceptual strategies.

Supplementary files, available at ANONYMOUS, give full data and details of the statistical analysis, for readers who wish to explore the data or try other analyses. The aim is reproducibility (see Stodden, Leisch & Peng 2014 for an overview; Zuraw 2015 for a recent example in linguistics): readers can re-run and then alter our statistical code.

## 2 Tongan background

Tongan is an Austronesian language, in the Polynesian subfamily, from the Kingdom of Tonga, with about 126,000 speakers (Lewis 2009). Tongan speakers have had extensive contact with British English, through missionary activity that became vigorous starting in the 1820s, and Tonga's status as a British protectorate for most of the 20th century (Lal & Fortune 2000: 141-142, 614-615). The Tongan lexicon is rich in loans from English.

The Tongan phoneme inventory is shown in (1).

(1) Tongan phoneme inventory

| | | | | | | |
|---|---|---|---|---|---|---|
| p | t | k | ʔ | i | | u |
| f | s | | h | e | | o |
| v | | | | | a | |
| m | n | ŋ | | | | |
| | l | | | | | |

[1] See Hall (2013) for a typology of marginal contrasts. Our analysis of vowel deletion fits into Hall's 'variation' category, but secondary stress does not quite fit into the typology.

The phoneme inventory plays little role in this paper, but being aware of the following common adaptations will make loans' sources easier to recognise: all English obstruents except [v] become devoiced, English [ɹ] becomes /l/, and various English vowel distinctions are collapsed.

Section 5 deals with vowel length, which is analyzed and transcribed two different ways in the literature. Tongan can be analyzed as having either a contrast between short and long vowels (Churchward 1953, Poser 1985), or sequences of vowels, in separate syllables, that sometimes happen to be identical (Feldman 1978, Schütz 2001, Taumoefolau 2002, Anderson & Otsuka 2006). The examples in (2) show transcriptions under each view. Transcriptions differ when stress falls on the first half of the vowel (sequence); [áː] vs. [á.a]; when stress falls on the second half, both views treat the sequence as disyllabic: a.á.

(2) Example words transcribed under two different views of syllabification

| *long vowels* | *V.V sequences* | |
|---|---|---|
| ku.m**áː** | ku.m**á.a** | 'rat' |
| m**àː**.l**óː** | m**à.a**.l**ó.o** | 'thank you' |
| m**a.á**.ma | m**a.á**.ma | 'light' |
| m**àː**.má-.ni | m**à.a**.má-.ni | 'this light' |
| h**úː** | h**ú.u** | 'enter' |
| h**u.ú**-.fi | h**u.ú**-.fi | 'open officially' |
| h**u.úː** | h**u.ú-.u** | 'enter-*definite*' |

In this paper we mostly use the 'V.V' transcription style (usually omitting the period marking the syllable boundary).

Tongan phonotactics are strict: in careful speech, every consonant must be followed by a vowel. Word shapes like CVCV, VCV, CVV are common, but not *CVCV**C**, ***C**CVCV, or *CV**C**CV. When an English source word has a consonant not followed by a vowel, that consonant must be deleted or, more commonly, have a vowel inserted after it. For example, English [æ**mb**ǽsədə][2] 'ambassador' becomes Tongan /ʔa**mip**asitoa/, with a vowel inserted after the [m]. All five Tongan vowels can serve as epenthetic; [i] is the most frequent, and an epenthetic vowel is often a copy of a nearby vowel. These inserted vowels play an important role in all three cases examined in this paper.

Our consultants were two women and one man, all from the Nuku'alofa area, who had been living in the United States for many years but speaking Tongan on a daily basis. Taumoefolau (1998) gives an overview of the role of English in Tonga over the past few decades. There are many monolingual Tongan speakers in Tonga, but English has become important in education and official writings, with increasing naturalistic contact through trade, tourism, migration, and media (125-136).

---

[2] All English pronunciations are based on the online Oxford English Dictionary (www.oed.com), with some adjustments.

## 3 Secondary stress

Stress is predictable in Tongan native words, but in loans, we will show that secondary stress is unpredictable from segments alone, and sensitive to whether a vowel is epenthetic.

### 3.1 Stress as previously described

Primary stress in Tongan is straightforward, falling on the second-to-last mora of a domain that includes stems, suffixes, and some enclitics (Churchward 1953: 4-5). That is, there is a moraic trochee at the end of each phonological word (Anderson & Otsuka 2006). The words in (3) illustrate these generalizations. To make examples easier to absorb, we include parentheses to indicate foot boundaries.

(3) Examples of primary stress

| | | | |
|---|---|---|---|
| (fále) | 'house' | fa(lé-ni) | 'this house' |
| fe(túʔu) | 'star' | fetu(ʔú-a) | 'starry' |
| (móhe) | 'to sleep' | mo(hé-ŋa) | 'bed' |
| (íka) | 'fish' | i(ká-a) | 'abounding in fish' |

Secondary stress has also been described (Churchward 1953, Feldman 1978, Schütz 2001). Feldman states that 'secondary stress falls on every second vowel from the end of each morpheme' (134), unless the next vowel has primary stress (e.g., in *fàka-aŋá-hi* 'criticise' the final vowel of the root *aŋa* receives primary stress because it is penultimate in the word, and therefore the penultimate vowel of *aŋa* does not receive secondary stress). Feldman's examples illustrate ways that his generalization cashes out:

(4) Secondary stress as described by Feldman

a) A morpheme of 4 or 6 moras has secondary stresses on alternating vowels
(kàa)(táki) 'sorry' (màa)(lòo)(lóo) 'rest'

b) The first part of a compound has penultimate secondary stress
**(mèʔa)**-va(ʔíŋa) 'toy' (thing+play)

c) A reduplicant has penultimate secondary stress
**(kàu)**-ka(ú-ʔi) 'wash' (REDUP+wash+transitive)

d) A disyllabic prefix has penultimate secondary stress
**(fàka)**-(ʔòfo)-(ʔófa) 'beautiful' (causative+REDUP+beautiful)

e) When there is a disyllabic suffix, the stem has penultimate stress
**ma(àma)**-(ʔáŋa) 'source of light' (light+location/source)

Long monomorphemes are rare and often contain reduplication-like repetitions or strings that exist as affixes in the language (*taki* and *maa*, in (4)a, are affixes that form verbs or adjectives with unpredictable meaning change, Churchward 1953: 243-244, 259-260).

Native Tongan vocabulary is silent on secondary-foot alignment: in hypothetical 5-mora */tatatatata/*, should there be an 'initial dactyl' (Prince 1983, Hayes 1995), as in *[(tàta)ta(táta)]*, or should all feet be aligned to the right, as in *[ta(tàta)(táta)]*? We failed to find convincingly monomorphemic native words of this shape (or with 7 moras). Because of epenthesis, monomorphemic loans from English offer many 5-mora cases and longer, and as we will see, both stress patterns occur.

Our purpose in this section is to investigate what determines whether an English loan is stressed as *[(tàta)ta(táta)]* or *[ta(tàta)(táta)]*. We show that secondary stress avoids falling on an epenthetic vowel.

### 3.2 Materials and methods

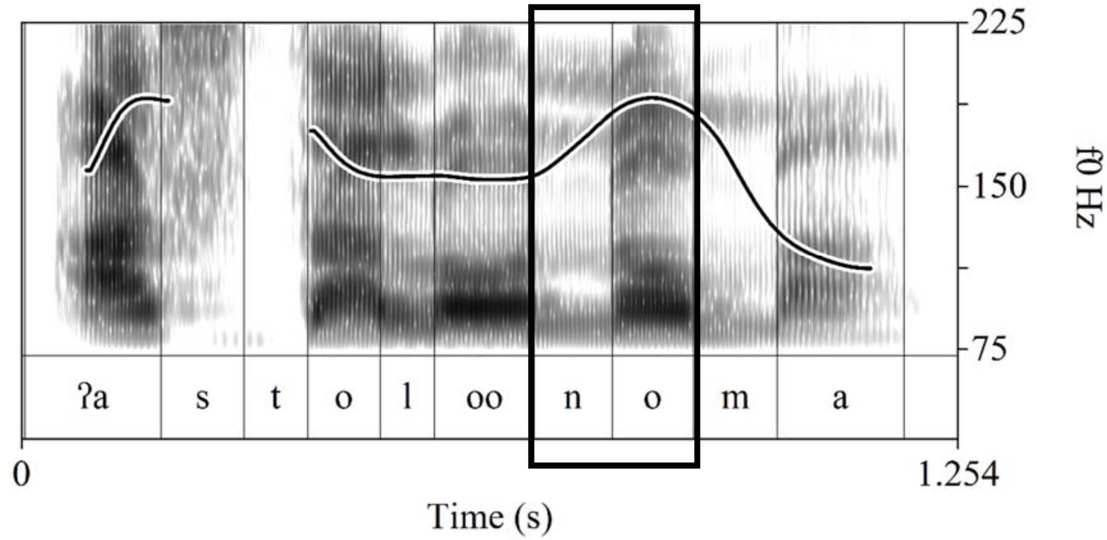
Churchward's dictionary (Churchward 1959) marks loans, and some additional words that Churchward and Tongan-speaking collaborator Feleti Vī thought the language might need to adopt, such as *palafini* 'paraffin'. We entered all these loans and suggested loans into a spreadsheet.

In each session with a consultant, we went through a few dozen words to check that the word was familiar, find out the consultant's preferred spelling, and check the gloss. Words deemed familiar were then recorded in a quiet room using Praat (Boersma & Weenink 2006) and a Logitech USB desktop microphone. Words were recorded in citation form, with two repetitions each. If both tokens were untranscribable because of hesitation, too much creaky voicing, or other problems, that word was re-recorded in a subsequent session. We recorded about 500 loans in total; after various exclusions discussed below, 134 were used to investigate secondary stress.

To objectively transcribe secondary stress, we first considered the realization of primary stress. Vicenik and Kuo (2010), working with the same consultants as we did, found that the primary-stressed syllable was usually associated with a pitch rise that begins around the beginning of the syllable and peaks around the end. We see this in the following token of [(ʔàs)toloo(nóma)] 'astronomer' from Speaker 2, where pitch rises over the primary-stressed syllable [no]:[3]

[3] All pitch tracks, spectrograms, and waveforms were made using Praat. Spectrograms are shown for reference; frequency range is 0 to 5000 Hz.

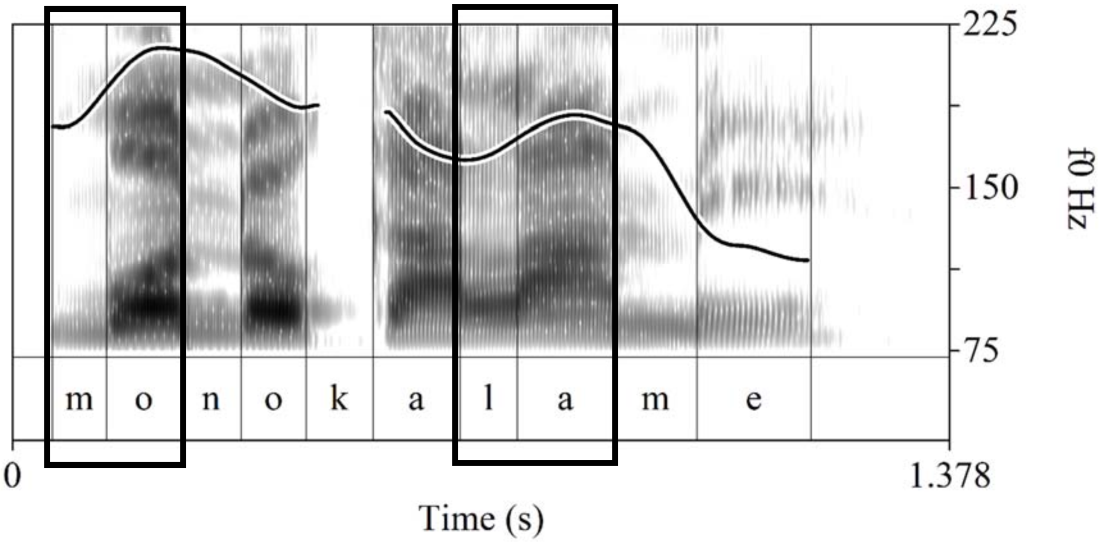
(5) Pitch track of [(ʔàs)toloo(nóma)] ‘astronomer’, primary stress boxed



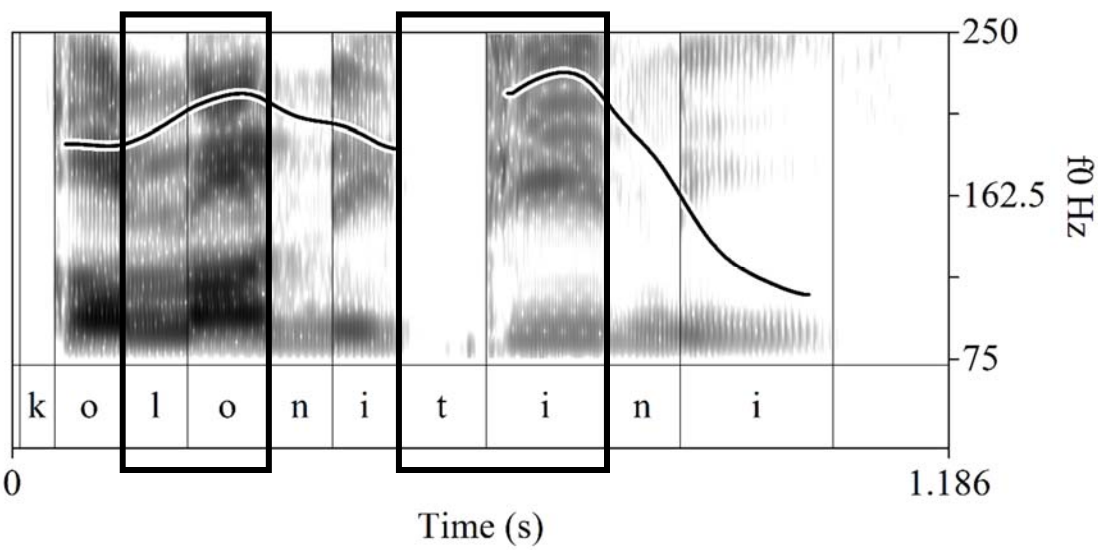

Preliminary investigation found that on syllables where we heard secondary stress, there was usually a similar pitch rise, as can be seen at the beginning of ‘astronomer’. (This token is also typical in having no clear sign of a tertiary stress, even though there is room for another foot in the middle of the word. Clear cues to tertiary stresses were rare, and we did not attempt to analyze it.) Therefore, we took rising pitch, followed by a fall, to indicate secondary stress on the syllable with the rising pitch. Using citation forms, rather than frame sentences, maximised the chance that secondary stress would display this cue. Doubtless there are other cues, but they are weaker and less reliable: Garellek & White 2015 found that fundamental frequency, RMS energy, duration, first formant, and voice quality were all reliably different between primary-stressed and unstressed vowels in Tongan. But looking at secondary stress versus no stress (in 4-mora words, where the position of secondary stress is uncontroversial), they found that only fundamental frequency was significantly different for all five vowels.

In (6) we see secondary stress on the first syllable: pitch rises on [mo], then falls on [no]. (The overall pitch of this token declines, so that the secondary stress’s f0 peak in [mo] is higher than the primary stress’s in [la].) In (7), by contrast, we see secondary stress on the second syllable: pitch rises on [lo], then falls on [ni].

(6) Pitch track of [(mòno)ka(láme)] ‘monogram’ from Speaker 1, secondary and primary stresses boxed




(7) Pitch track of [ko(lòni)(tíni)] ‘quarantine’ from Speaker 2




The three authors independently transcribed each token. There were tokens with a flat pitch contour over the first two syllables and thus no discernable secondary stress according to our criteria. Such tokens were not used in the analysis below. There were also tokens where one of the first two vowels was deleted, such as the one shown in [ʔàs_toloonoma] in (5) above, from /ʔasitoloonoma /; we coded the deleted vowel as unstressed, but the coding of the other vowel’s stress still depended on whether it bore a pitch rise. In our original set of 306 tokens that meet the further criteria set out below, we disagreed on 19 (6%), in almost every case because one or two transcribers detected no clear secondary stress on either of the first two vowels, while the rest did detect some. There were only 3 cases where we transcribed opposite secondary stress; these were excluded. For the remaining 16 items, we retranscribed, and went with majority rule if there was still disagreement.[4]

Once we had our speakers' pronunciations of all the words, we excluded certain word shapes. (Because speakers often disagreed slightly with the dictionary form, we would not want to exclude items a priori based on their dictionary form.) We included only words of the form #(C)$V_1$(C)$V_2$(C)$V_3$...$\acute{V}$(C)V#, such as *ʔolokaholo* 'alcohol'. Requiring at least one vowel between $V_2$ and the main stress means that there is enough room before the primary stress for either $V_1$ or $V_2$ to be stressed: actual [(ʔòlo)ka(hólo)] and hypothetical [ʔo(lòka)(hólo)] are both well footed. In shorter words like *ʔamelali* 'admiral', there is only one place for a non-final foot, [(ʔàme)(láli)]; such words tells us nothing about subtler footing preferences, and are excluded.

We placed additional restrictions on vowel sequences, because of additional stress generalizations we noticed. We often observed 'broken' vowels like [eé], with stress on the second half, in a word's last foot (*m**eé**si* 'mercy'), or the last foot of an obvious compound member (*h**aà**fe-kalaúni* 'half-crown'). But we did not observe broken vowels earlier in a word: we transcribed 43 tokens like [(ʔ**àa**)ke(tíka)] 'architect', and none like *[ʔ**a(à**ke)(tíka)]. Even when neither half of the sequence was stressed, it was rare for a foot boundary to split it: we observed 48 tokens like [pa(l**òo**)(mési)] 'promise' and just 2 like [(sèv**a)a**(níti)] 'servant'; 53 tokens like [(ʔèle)(f**aa**)(níte)] (with no clear tertiary stress, but at least the possibility that the two *a*s are footed together) and none like *[ʔe(lèf**a**)**a(**níte)]. There must be a penalty against […a(á…] and […a)a…].[5] Therefore, if any two of a word's first four vowels were identical and immediately adjacent, we excluded the word (*ʔ**aa**ketika*, *pal**oo**mesi*, *elef**aa**nite*).

For similar reasons, we did not allow two adjacent vowels (in the word's first four vowels) to be of rising height; that is, *ai*, *ae*, *au*, *ao*, *ei*, *eu*, *oi*, or *ou* (*t**ai**namiki* 'dynamics', *kal**ai**meti* 'climate', *ʔepon**ai**te* 'ebonite' were excluded). This is because sequences of rising height may have special stress behavior too. Churchward (1953: 4) claims that in a word like *tauhi* 'to keep', 'stress may fall either on the *u* or on *au* as a whole', and similarly for words with *ao*, *ai*, or *ei*; Schütz (2001: 319-321) suggests that the variation for these 'potential diphthongs' is conditioned by phrase position and speech style and rate. Our transcription of primary stress in words ending with sequences like …auCV was often uncertain. Garellek and White (2010) find acoustic evidence that VV sequences with rising height tend to belong to a single syllable, while VV sequences with falling height tend to belong to two separate syllables. We might therefore expect that *[t**a(ì**na)(míki)] is ill-formed because it splits an *ai* sequence into separate syllables. This prediction was also borne out in our data: we found 36 tokens like [(t**ài**)na(míki)] and only 6 like [p**o(ì**ni)se(tía)] 'poinsettia'; 7 like [ka(l**ài**)(méti)] and 4 like [(tìs**a)i**(péle)] 'disciple'; 24 like [(ʔèpo)n**ái**te] or [(ʔèpo)n**aí**te] (leaving open where the main-stress foot boundaries would be) and only 1 like [si(kàla**)(í**pe)] 'scribe'.

---

[4] Two items were transcribed by only two researchers, but given our low rate of disagreement this wasn't worrying.

[5] This constraint might require adjacent, identical vowels to be in the same foot. Or, if all such sequences are really underlying long vowels (going against Prince & Smolensky's 1993 Richness of the Base), it might be a constraint requiring long vowels to be stressed on their first mora. Whatever the constraint is, it is outranked by the requirement for a word to end with a bimoraic, trochaic foot, as in [me(é.si)] 'mercy'. Thus, 'breaking' (Poser 1985, Mester 1992) occurs only at the end of a word.

Occasionally, a speaker's two repetitions of a word differed in meeting the criteria above, such as [(sèe)keli(táli)], [(sèke)li(táli)] 'secretary', where only the second pronunciation meets our criteria. We take it that there is variation or uncertainty in the underlying form, and use the repetition that meets our criteria. There were also five cases, which we simply excluded, whose repetitions differed purely in the location of secondary stress, such as [(kòni)sinaa(níte)] 'consonant', [ko(nìsi)naa(níte)].

We excluded four additional categories. First, we excluded probable compounds, such as /simolo-pookisi/ 'small pox', because they should contain two separate stress domains. Second, we excluded a few religious words, such as *ʔevaŋelioo* 'evangel, gospel', that could not be straightforward adaptations from English (what would be the source of the final *oo*?). The source might be another language, perhaps as pronounced by English-speaking missionaries, so that we were unsure of the source stress. Third, we excluded two tokens produced with exceptional antepenultimate stress, because there were not enough vowels before the primary stress to allow both options for secondary stress: [sèmeneelío] ~ [sèmenélio] 'seminary' (excluded anyway because the final [o] makes the source obscure), and [ʔèlevásio] 'elevation'.[6] Fourth, we excluded words that begin with /i/ or /u/ and another vowel, because they tend to become glides: /uesiliana/ [(wèsi)li(ána)] 'Wesleyan'; we also excluded one word with a high vowel between two other vowels, /noauee/ 'Norway', because after gliding there were only four surface vowels: [(nòa)(wée)].

### 3.3 Results

All this culling left us with 294 tokens from the three speakers, covering 134 types. Of these, 189 (64%) have stress on $V_1$ and 105 (36%) have stress on $V_2$.

$V_1$'s or $V_2$'s correspondent in the English source can be a stressed vowel, an unstressed vowel, or no vowel at all. We want to know if this affects where secondary stress falls.

Raw results are shown in (8). Columns divide the data according to $V_1$'s status, and rows according to $V_2$'s. The percentage of words in each category with secondary stress on $V_1$ is shown; remaining words have stress on $V_2$. An example word is given for each stress pattern. Unfortunately, some of the cells are empty—for example, there were no qualifying items where both $V_1$ and $V_2$ are epenthetic—but we can still make some direct comparisons.

(8) Percentage of items stressed on $V_1$

| | $V_1$ epenthetic | $V_1$ present but unstressed in English | $V_1$ stressed in English |
|---|---|---|---|
| $V_2$ epenthetic | | 13/15=87%<br>ʔ**à**sitolòonóma 'as_trónomer'<br>*vs.*<br>ʔin**ì**sipeekíta 'ìn_spéctor' | 79/95=83%<br>k**à**ŋikalúu 'kàn_garóo'<br>*vs.*<br>ʔep**è**likóti 'ápricot' |
| $V_2$ present but | 0/1 | | 49/71=69%<br>m**ò**nokaláme 'mónogràm' |

[6] A third token from the same speaker, [kònteneeníti] ~ [kòntinénitḭ] 'continent', was included because it was long enough.

| | | | |
|---|---|---|---|
| unstressed in English | (no examples of stress on $V_1$)<br>pol**ò**minísi ‘b_lancmánge’<br>[bləmɒ́nʒ] | | *vs.*<br>hip**ò**potamási ‘hìppopótamus’ |
| $V_2$ stressed in English | 17/65=26%<br>p**à**lopaléma ‘p_róblem’<br>*vs.*<br>fal**à**kiséni ‘f_ráction’ | 31/47=66%<br>tèmokalasi ‘demócracy’<br>*vs.*<br>ʔap**è**nitíki ‘appéndix’ | |

Comparing the cells in the bottom row, where $V_2$ corresponds to a stressed English vowel, we see that it matters whether $V_1$ is epenthetic (words like /f**a**lakiseni/ ‘f_ráction/) or corresponds to an unstressed English vowel (/ʔ**a**penitiki/ ‘**a**ppéndix’). Epenthetic $V_1$s are stressed only 26% of the time, as compared to 66% for non-epenthetic $V_1$s.

In the last two columns we can also make direct comparisons along the $V_2$ dimension. There is a higher rate of stressing $V_1$ when $V_2$ is epenthetic (words like /ʔasitoloonoma/ ‘astronomer’ and /kaŋikaluu/ ‘kangaroo’, 87% and 83%; overall 84%) than when it is not (words like /ʔapenitiki/ ‘appendix’ and /monokalame/ ‘monogram’, 66% and 69%).

Are these differences significant? The cells of the table may be unbalanced with respect to which speaker the data come from, and the qualities of $V_1$ and $V_2$ (/a,e,i,o,u/). Therefore, we used regression to control for these factors.

The logistic regression model, shown in (9), was fitted using the *bayesGLM()* function of the *arm* package (Gelman et al. 2010) in R (R Core Team 2014). This function has the advantage of handling the ‘separation’ found in the data (cases where some combination of independent variables completely predicts the outcome, such as $V_1$ status = epenthetic and $V_2$ status = unstressed).

The dependent variable is whether stress falls on $V_1$ (as opposed to $V_2$); variables that promote stress on $V_1$ have positive coefficients and those that promote stress on $V_2$ have negative coefficients. The independent variable of interest encodes the status of the first two vowels: rather than treat the rows and columns of (8) as two separate variables, with a lack of data for certain combinations of values, we use a single variable ‘V status’, with values like ‘$V_1$ is epenthetic, $V_2$ corresponds to stressed’. We excluded ‘blancmange’ to avoid having only one token whose value of this variable is ‘$V_1$ is epenthetic, $V_2$ corresponds to unstressed’.

Four independent variables were included to control for possible skews in the data. First is *speaker*, in case speakers have different baseline rates of stressing $V_1$.[7] Second is a binary variable encoding whether the word begins with a possible English cluster. The idea is that in a word like /kolonitini/ ‘quarantine’, even though the first vowel is not epenthetic, speakers might have formed a generalization, based on other loans, that a vowel in the environment #k__l tends to be epenthetic,

[7] With only three levels, we did not treat speaker as a random effect. See literature review in Hox, Moerbeek & Schoot (2010: 46) and recommendation in Snijders & Bosker (2011 :48): the minimum number of recommended levels for a random effect varies, but is always higher than three.

and thus might treat it differently. If the first two consonants in the word were /p__l, f__l, t__l,[8] k__l, s__l, s__m, s__n, s__p, s__t, s__k/, this variable was coded as 'yes'. Third and fourth are vowel quality for $V_1$ and $V_2$, each with five levels: /a [reference level], e, o, i, u/.

We used R's *step()* function for model selection, which tries adding and omitting various factors, seeking improvement in the Akaike Information Criterion. Starting from a model with all the variables listed above, plus an interaction between speaker and every other variable, the best model found is shown in (9):

(9) Logistic regression model for secondary stress. Dependent variable is whether stress falls on $V_1$. Independent variable of interest is in bold.

Coefficients:

| | Estimate | Std. Error | z value | Pr(>\|z\|) | |
|---|---|---|---|---|---|
| (Intercept) | -1.110 | 0.371 | -2.99 | 0.003 | * |
| speaker | | | | | |
| = Speaker 1 | *reference level* | | | | |
| = Speaker 2 | 0.160 | 0.341 | 0.47 | 0.638 | |
| = Speaker 3 | 0.834 | 0.347 | 2.40 | 0.016 | * |
| V1 | | | | | |
| = a | *reference level* | | | | |
| = e | 0.903 | 0.440 | 2.05 | 0.040 | * |
| = o | -0.479 | 0.416 | -1.15 | 0.250 | |
| = i | -0.860 | 0.364 | -2.36 | 0.018 | * |
| = u | 0.122 | 1.092 | 0.11 | 0.911 | |
| **V status** | | | | | |
| **= V1 epenthetic, V2 corr. to stressed** | ***reference level*** | | | | |
| **= V1 corr. to stressed, V2 epenthetic** | **2.397** | **0.419** | **5.73** | **< 0.001** | ***** |
| **= V1 corr. to unstressed, V2 epenthetic** | **2.597** | **0.740** | **3.51** | **< 0.001** | ***** |
| **= V1 corr. to unstressed, V2 corr. to stressed** | **1.635** | **0.443** | **3.69** | **< 0.001** | ***** |
| **= V1 corr to stressed, V2 corr. to unstressed** | **1.656** | **0.415** | **3.99** | **< 0.001** | ***** |

The intercept represents the predictions for the default (reference) level of all independent variables, and the coefficients for the other factors are added to the intercept when relevant. For an example like *kasitomaa* 'customer' uttered by Speaker 1, *speaker* and *$V_1$* have their default values, and *V status* has the value '$V_1$ corresponds to stress, $V_2$ is epenthetic', so we add the intercept, -1.110, and the coefficient 2.397. The model's predicted rate of stressing $V_1$ is $1/(1+e^{-(-1.110+2.397)})$ = 78% (actual rate for the 13 such words: 69%).

Stepping through the coefficients, we see that Speaker 3 stresses $V_1$ significantly more often than Speaker 1 (the reference level). Quality of $V_1$ has some effect: /i/ is stressed significantly less often than /a/, and /e/ significantly more often than /a/.[9] Turning to the factor of interest, *V status*, we see that every other value has a higher rate of stressing $V_1$ than does the reference level,

[8] English [ɹ] is adapted as Tongan [l], as in *talanisimita* 'transmitter'. Thus, at least some Tongan loan that begin *tVl*... come from an English word that begins with a consonant cluster.

[9] It is surprising to see a stressability hierarchy of e > a > i, given typological findings that low vowels should be more stressable than mid vowels (e.g., Kenstowicz 1997, de Lacy 2007).

where $V_1$ is epenthetic and $V_2$ corresponds to a stressed English vowel. (Using the function *glht()* from the *multcomp* package, Hothorn, Bretz & Westfall 2008, we find that none of the other pairwise differences between levels of this variable are significant.) In terms of minimal comparisons within a row or column of the table in (8), where only one vowel's status is changed at a time, this means there is a significant difference between words like 'p_róblem' and words like 'potássium'—that is, a significant difference between an epenthetic and non-epenthetic $V_1$.

### 3.4 Discussion

The secondary stress data we have just seen indicate that Tongan speakers are sensitive to the English contrast between #CC and #CVC, such as '**pr**edicate' (***pal****ètikási*) vs. '**par**agraph' (***pàl****akaláfi*), even though there is no such contrast among native words: at least before contact with English, speakers did not have experience perceiving near-minimal pairs analogous to English *please* and *police*.

Tongan has much in common with Japanese, where perception of the CC/CVC difference has been studied in detail. Like Japanese, Tongan has extensive vowel devoicing and even some deletion (see section 4), so that an English CC sequence might actually be close to a Tongan /...CVC.../ sequence that has undergone vowel deletion. In other words, the Tongan 'V' category in the context C__C includes a range from Ø to V, and crucially does not contrast with Ø. This should make perceiving the difference between Ø and V even more challenging in this environment: if Ø could be perceived as an extremely poor exemplar of V, the difference might be easier to perceive, but if Ø is actually a reasonably common exemplar of V, then we expect the perceptual space to be warped so that Ø and V are perceived as more similar (e.g., Iverson & Kuhl 1996, where English speakers perceived the distance between two good tokens of /l/ as smaller than the distance between a good and a bad token; and similar results for English /i/ in Iverson & Kuhl 1995).

As mentioned in the introduction, the Japanese speakers studied by Dupoux, Kakehi, et al. (1999), even more-advanced learners of English or French included in one experiment, had difficulty distinguishing CC and CVC, though they performed above chance in an ABX task. Peperkamp and Dupoux (2003) suggest that loan adaptation is difficult enough that perceptually driven nativizations like vowel insertion will happen consistently.

By contrast, Kubozono (2006) argues that English CC and CVC are treated in subtly different ways in Japanese loan adaptation, and this is what we seem to see in Tongan.

## 4 Vowel deletion

Previous authors have noted that Tongan vowels sometimes devoice or even delete. Feldman (1978) states that for native or other Polynesian words 'in all but the very most careful style of speaking' (137), high vowels are devoiced if short, unstressed, preceded by a voiceless consonant, and either utterance-final ([tápi̥] 'wipe messily') or morpheme-final before a voiceless consonant ([fàka-fòki̥-fáa] 'suddenly'. Feldman reports that /a/ is devoiced under the same conditions but only after /h/, with the result often being [xː] ([táhḁ] ~ [taxː] 'one'). (Morton 1962 gives somewhat different generalizations, including for mid vowels.)

Feldman also gives examples of loanwords that devoice or delete in additional environments, such as [pàsi̥kála] ‘bicycle’ (138), whose /i/ shouldn’t devoice because it’s not followed by a word or morpheme boundary, or [pènsimán] ‘Benjamin’, spelled *Pen**i**siman**i***, where the preceding consonant is voiced /n/. Feldman proposes that the devoicing and deletion happen ‘presumably to bring the Tongan pronunciation into closer conformity with the English’ (138).

Although Feldman’s description limits vowel *deletion* (as opposed to mere devoicing) to loanwords, we found that it was fairly common in native words, even where not predicted by Feldman’s rules. The following examples are from Sailor (2010), working with the same consultants as we did. (Sailor finds, impressionistically, that deletion and devoicing are more common in loans than in native words.)

(10) Vowel deletion examples from Sailor 2010

| | | | |
|---|---|---|---|
| /u/ | [tuʔ_] | ‘stop’ | |
| /a/ | [f_ka-maʔa-maʔa] | ‘on sale’ | *not followed by boundary* |
| /i/ | [poŋ_-poŋi] | ‘morning’ | *preceding C is voiced* |

Vowel deletion was also common in our corpus of loan pronunciations. That is, there were vowels that were not pronounced, but that the consultant included in the spelling of a word, and that sometimes were pronounced in other tokens of the same word. As Sailor points out, sometimes this leads to a pronunciation closer to the English—as in /ʔasipesitosi/ [ʔàs_pes_tós_] ‘asbestos’ or /ʔapenitiki/ [ʔapèn_tíki̥] ‘appendix’, where the deleted vowels had no correspondent in the English form—but not always, as in /ʔamipasitoa/ [ʔàmipas_tóa] ‘ambassador’, where the deleted vowel corresponds to a real vowel in the English form, and the other /i/ in the word, despite being epenthetic, does not delete.

In this section we show that the vowel /i/ is more likely to delete if it has no English correspondent.

### 4.1 Methods

The items for this analysis were taken from the same recordings as the items in the secondary stress analysis, but are a different (overlapping) subset. Nearly all qualifying deletions (criteria given immediately below) were of /i/, so we limited our analysis to that vowel (there were five qualifying deletions of /e/, and none of other vowels). We chose items with at least one unstressed /i/, according to the consultant’s spelling. To make sure that the /i/ was unstressed, the preceding or following vowel, or both, had to have been transcribed as stressed. Because final /i/s are almost all epenthetic, we excluded them. We also excluded /i/s in the first syllable, because whether or not they are epenthetic is strongly confounded with the surrounding consonants; for example, if the surrounding consonants are s__k, s__t, or s__p, the /i/ is almost always epenthetic. Therefore, we used only words with the following patterns:

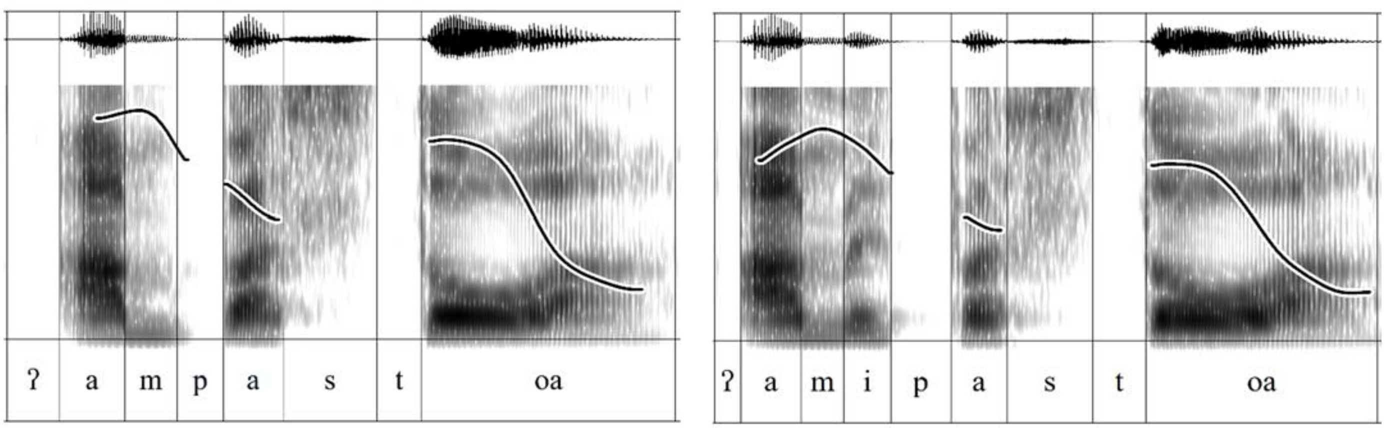
(11) Tokens selected for vowel-deletion analysis

| *pattern* | *# of tokens* | *examples* |
| --- | --- | --- |
| between two stresses | 159 | ʔàn**i**téma ‘an_them’, ʔopèl**i**káto ‘obbl**i**gato’ |
| after a stress | 93 | ʔòs**i**talési ‘os_trich’, sakìl**i**fisío ‘sacr**i**fice’[10] |
| before a stress | 41 | làven**i**táa ‘laven_der’ |

There were 22 items with two qualifying /i/s, such as *sakìl**i**f**i**sío* ‘sacrifice’.

Looking at the waveforms and spectrograms of these items, there was a clear difference in the recordings between deleted and non-deleted /i/s, so only one author did the additional codings needed for this section. An example of the difference between deleted and non-deleted /i/ is shown in (12), for two tokens of *amipasitoa* ‘ambassador’ produced in immediate succession.

(12) Two tokens of *amipasitoa* ‘ambassador’ by Speaker 2 (f0 range from 75 to 225)




Although word-final devoiced /i/ was common, devoiced /i/ was rare in our target contexts,[11] and we grouped it with deletion, in line with the analysis presented below in 6.1. If deletion occurred in only one repetitions, we coded the item as displaying deletion, on the idea that what is of interest is whether a vowel *can* delete (there were 9 such items; the supplementary materials include versions of the regression model below in which these items are instead excluded or treated as non-deleted).

### 4.2 Results and discussion

From the raw data in (13), it is clear that epenthetic /i/ is more likely to delete, although the three speakers have different global rates of deletion. Unfortunately, there are not enough items where unstressed, word-medial /i/ corresponds to a stressed vowel in English to see whether stress in English matters.

[10] The final [o] suggests a non-English origin. In this section we allowed such words, because the important factor, epenthesis status, was clear regardless of etymology.

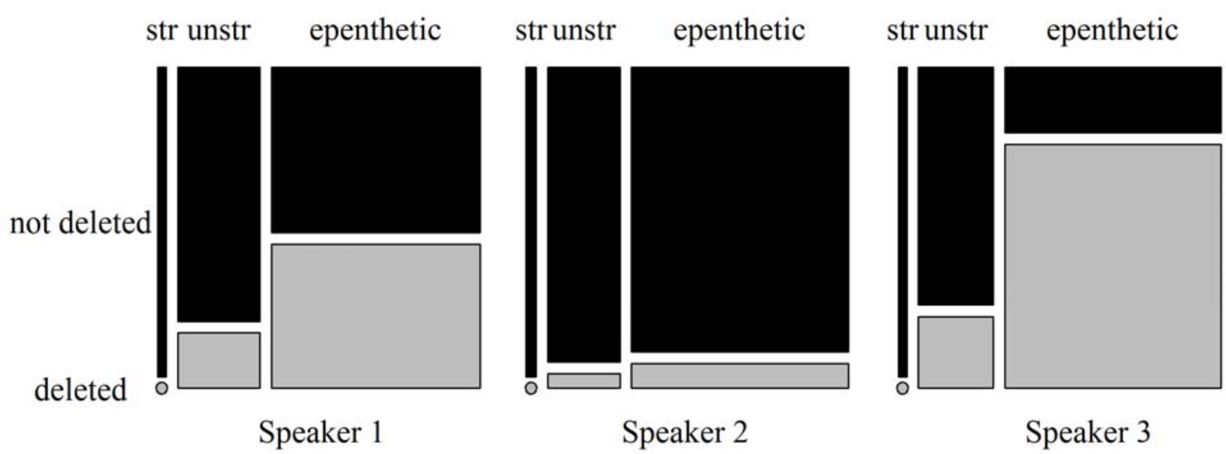
[11] There were three relevant tokens with (possible) devoiced vowels: one token of ‘as_bestos’ was transcribed by one coder with a devoiced /i/ and by the other two with no vowel at all; one token of ‘nas_turtium’ and one token of ‘monas_tery’ were transcribed as having a possible devoiced /i/ by two coders and with no vowel by the other coder.

(13) Percentage of /i/s that delete

| | /i/ corresponding to stressed English V | /i/ corresponding to unstressed English V | epenthetic /i/ |
|---|---|---|---|
| Speaker 1 | 0/3 | 5/28 = 18% | 33/71 = 46% |
| Speaker 2 | 0/3 | 1/21 = 5% | 5/63 = 8% |
| Speaker 3 | 0/3 | 6/26 = 23% | 59/75 = 79% |
| *overall* | *0/9* | *12/75 = 16%* | *97/209 = 46%* |
| *example with deletion* | *NA* | pàles_téni ‘prés**i**dent’ | kìukam_páa ‘cúcum_ber’ |
| *example without deletion* | mèl**i**tiáne ‘mer**í**dian’ | tipòos**i**tóa ‘depós**i**tor’ | tàlanis**i**míta ‘trans_mítter’ |

The mosaic plots in (14) show the same data. The grey regions represent /i/-deletion cases, and the black regions non-deletion. The width of each column (stressed, unstressed, epenthetic) reflects the proportion of cases for that speaker that come from that category. The circles at the bottoms of the ‘stressed’ columns for all three speakers indicate zero deleted /i/ in this category.

(14) Deletion (grey) and non-deletion (black) for each speaker




To test significance, we fitted a mixed-effects regression model, using the *glmer()* function in R’s *lme4* package (Bates et al. 2014), shown in (15). The dependent variable is whether /i/ is deleted. The independent variable of interest is whether Tongan /i/ is epenthetic or corresponds to an unstressed English vowel—we excluded the small number of cases where /i/ corresponds to a stressed English vowel. Speaker is treated as a fixed-effect independent variable with three levels. We included an independent variable coding whether the target vowel would be footed (as in [(ʔòs**i**)ta(lési)]), or not (as in [(làve)n**i**(táa)]). Finally, consonantal environment is treated as a random effect (not shown in (15)); in our data it has 36 levels, such as ‘s__t’ and ‘n__t’.

(15) Logistic regression model for vowel deletion. Consonantal environment (not shown) is random effect. Independent variable of interest is in bold.

| | Coefficient estimate | Std. Error | z value | p |
|---|---|---|---|---|
| (Intercept) | -2.927 | 0.715 | -4.09 | < 0.001 |
| speaker | | | | |
| = Speaker 1 | *reference level* | | | |
| = Speaker 2 | -2.727 | 0.530 | -5.15 | < 0.001 |
| = Speaker 3 | 1.543 | 0.391 | 3.95 | < 0.001 |
| is it footed? | | | | |
| = no | *reference level* | | | |
| = yes | 0.576 | 0.550 | 1.05 | 0.295 |
| **V status** | | | | |
| **= non-epenthetic** | ***reference level*** | | | |
| **= epenthetic** | **1.492** | **0.534** | **2.79** | **0.005 *** |

The intercept represents the case where the speaker is Speaker 1, the target vowel is not footed, and the target vowel is not epenthetic (in an average consonantal context). To get the model's predicted rate of /i/ deletion in that case, take $1/(1+e^{-(-2.927)}) = 5\%$. As we saw in (13), the speakers differ in their baseline rates of deletion, and this is significant in the model: Speaker 2 has a significantly lower rate of deletion than Speaker 1, and Speaker 3 has a significantly higher rate of deletion than Speaker 1.

As for our variable of interest, if the target vowel is epenthetic it has a significantly higher rate of deleting. The supplemental materials include comparison to an inferior model with a non-significant interaction between speaker and V status. The lack of significant interaction means there is no evidence that the three speakers had different degrees of sensitivity to the English pronunciation.

As in the case of secondary stress, Tongan speakers are showing sensitivity to whether a vowel in a loan is epenthetic. In the secondary stress case, we saw that Tongan speakers were distinguishing CC from CVC near the beginning of a word, though native Tongan words have no such contrast. Here, we see that they are also distinguishing CC from CVC further into the word.

## 5 Vowel length

### 5.1 Background and methods

Vowel length is contrastive in Tongan. That is, we see contrasts such as those shown in the left column of (16). As discussed in Section 1, some authors analyze these 'long vowels' as sequences of identical vowels (second column of (16)); under that analysis there is a contrast between, e.g., [a] and [aa], rather than [a] and [aː].

(16) Examples of vowel-length contrasts

| | | | |
|---|---|---|---|
| *V: notation* | *VV notation* | | |
| káka | kaka | 'to climb' | |
| kaká: | kakáa | 'parrot' | |
| kàːkáː | kàakáa | 'to cheat' | (Churchward 1959: 3) |
| | | | |
| káta | káta | 'to laugh' | |
| kaáta | kaáta | 'garter' (loan) | (Churchward 1959) |

English vowels can be adapted as short or long. Some of this probably derives from the vowel's duration in English. English stressed and/or tense vowels, being longer than unstressed or lax[12] vowels in English, should tend to be perceived by Tongan speakers as long. For example, the long vowel in [(kàŋi)ka(lúu)] 'kangaroo' is both stressed and tense in English ([kæ̀ŋgəɹ**ú**ː]).

But unstressed English schwa can be adapted as either long or short, as in [(hèli)kope(t**áa**)] 'helicopter', from English [hɛ́lɪkɒ̀pt**ə**], and [ta(làni)si(míta)] 'transmitter', from English [trɑ̀ːnsmít**ə**]. It seems unlikely that there is much length difference in the final vowels of the two English words. The important difference is, as we demonstrate in this section, that in [ta(làni)si(mít**a**)], the final short vowel causes primary stress to fall on the preceding syllable's /i/, which corresponds to English [í]; in 'helicopter', by contrast, a final short vowel, *[(hèli)ko(pét**a**)], would cause stress to fall on /e/, which has no English correspondent.

Schütz (1970) calls this tendency a 'method [for matching the English, which] doubles the final vowel. Although the penultimate vowel of this form does not correspond to the primary stressed vowel of the E[nglish] model, the stress has been moved from the inserted vowel. For example, E[nglish] /plǽstə/ would produce T[ongan] *pulasíta. The form that does occur is palasitaa.' Schütz notes that this phenomenon is not regular enough to write a rule for it.

Kenstowicz (2007) analyzes vowel length in Fijian loans from English (corpus from Schütz 1978) and finds something very similar. Fijian has a similar prosodic system to Tongan's, with contrastive length and a moraic trochee at the right edge of the word. Kenstowicz shows that length is deployed to achieve a better match between the prosody of the Fijian and the English forms, as listed in (17).

(17) Functions of vowel length in Fijian loans

- cause stress on an antepenult whose English correspondent is stressed
  (k**ò**ː)(lóni) 'c**ó**lony' vs. t**o**(váko) 't**o**bácco' (319)

- cause stress on an ultima whose English correspondent is stressed
  q**i**(táː) 'gu**i**tár' vs. (f**í**va) 'f**é**ver' (319)

- prevent stress on an epenthetic penult
  (sìsi)(táː) 'sist**er**' vs. pa(jáma) 'pajám**a**' (319, 324)

---

[12] E.g, Hillenbrand et al. (1995) found that American English /ʌ, ɪ, ɛ, ʊ/ were consistently shorter than other vowels.

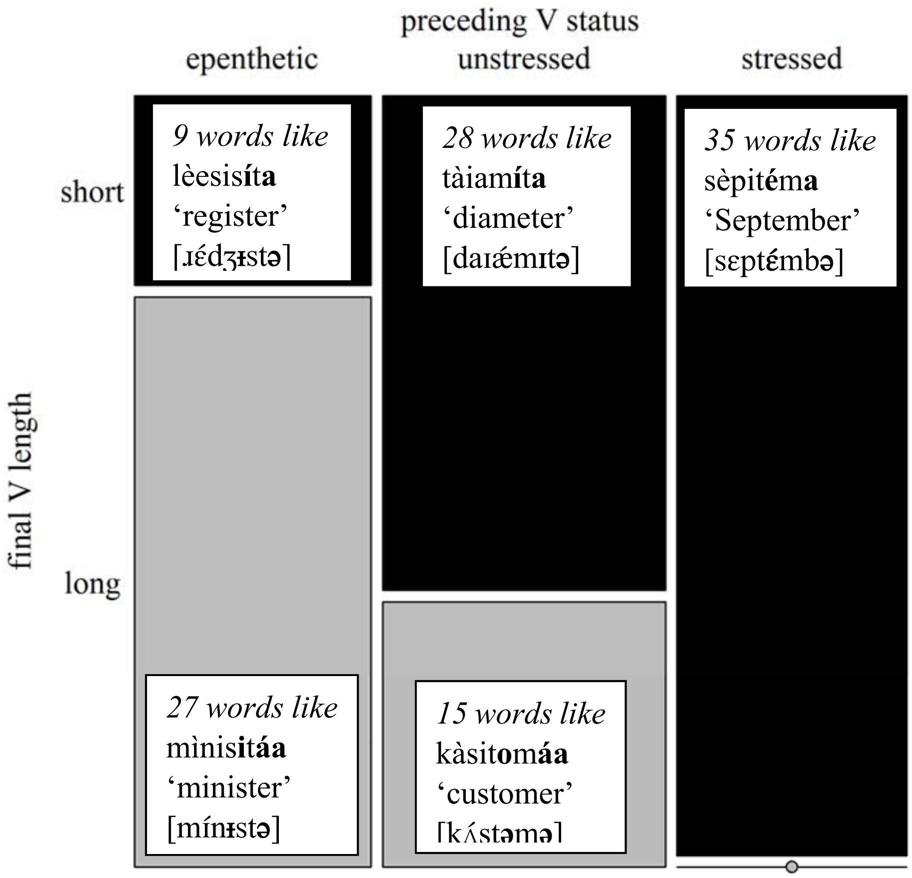
This section shows that Schütz's observation for Tongan epenthetic penults, and Kenstowicz's for Fijian epenthetic penults, also hold in our corpus of Tongan loans. We further show that final vowel length is sensitive to the English stressed/unstressed distinction: even when the penult is not epenthetic, final vowel length is more likely if the penult was unstressed in English than if it was stressed. Thus, the vowel length data provide (further) evidence that Tongan speakers perceive the difference between English CVC and CC, and the difference between English stressed and unstressed vowels.

We took from our recorded corpus all those words where the final Tongan vowel corresponds to an unstressed vowel in English, and analyzed the length of that final vowel.

### 5.2 Results

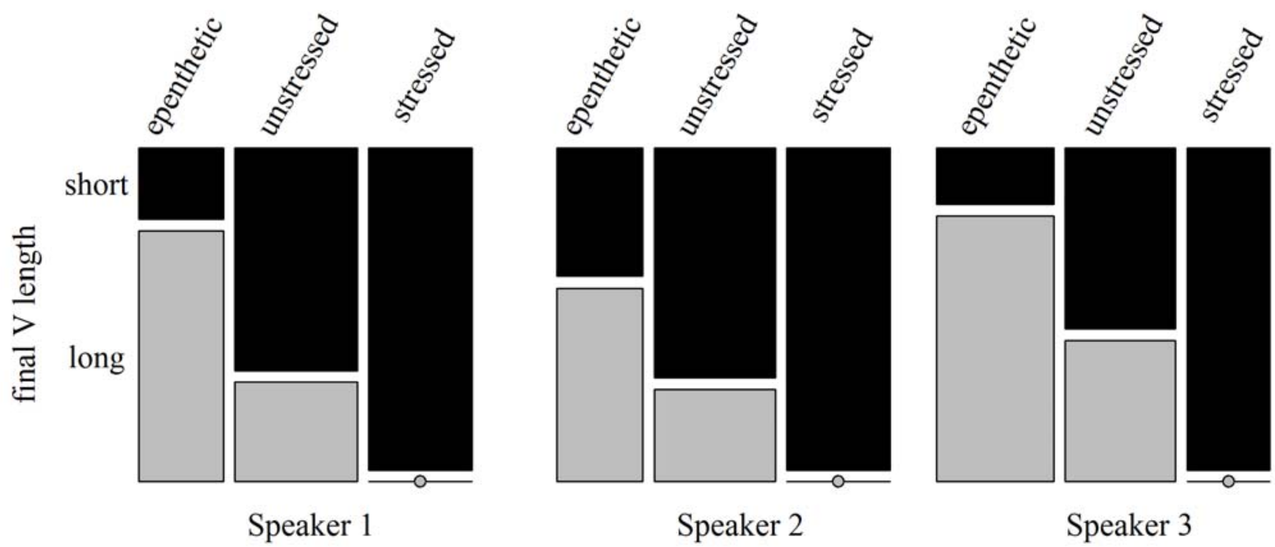
There were 114 relevant tokens. Looking at the raw numbers, we see a clear three-way distinction, as plotted in (18). If the preceding vowel has no correspondent in English (epenthetic), the final vowel was adapted as long 75% of the time. If the preceding vowel corresponds to an unstressed vowel it was adapted as long 35% of the time, and if it corresponds to a stressed vowel it was never adapted as long.

(18) Rate of adaptation with final long vowel for English words that end in an unstressed vowel (circle at bottom of third column represents token count of zero)



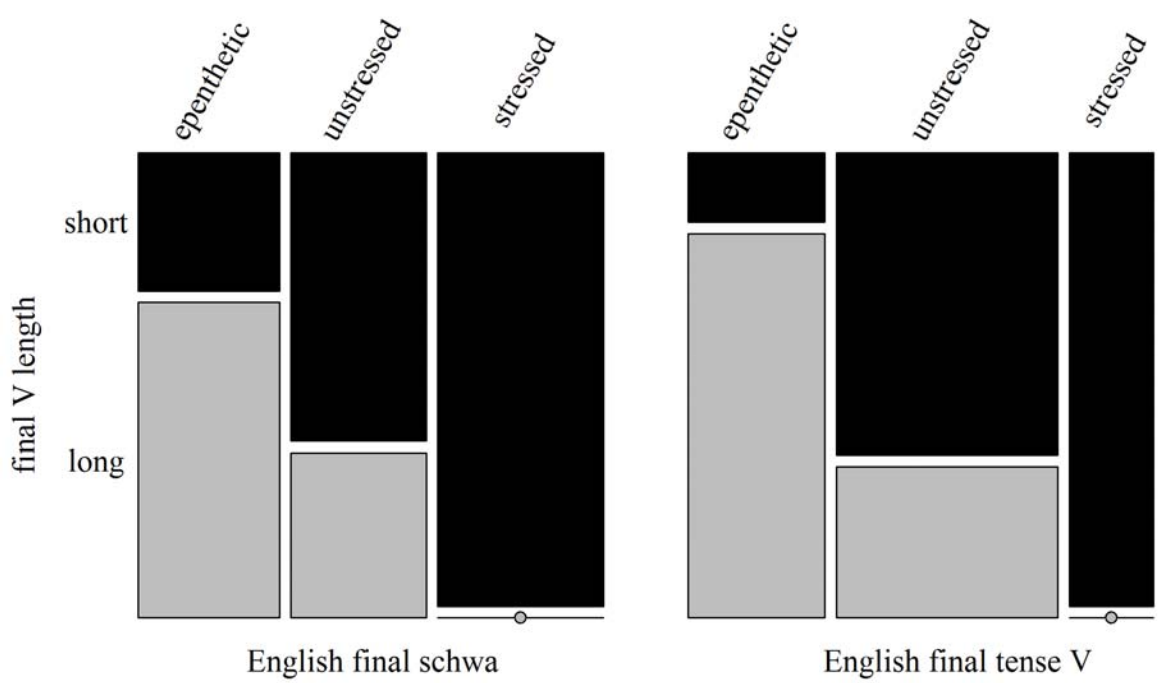

Results are similar for all three speakers, as shown in (19).

(19) Breakdown of (18) for each speaker




We can also see that the results are similar whether the final unstressed English vowel is [ə] (72 tokens) or a tense vowel ([oʊ] or [i], 42 tokens):

(20) Breakdown of (18) for each vowel type




To test the significance of the penult's status, we used a logistic regression model. Because there are no random effects in this case, we were once again able to use the function *bayesGLM()*, which can handle the zero rates of lengthening for stressed English penults. The binary dependent variable is whether the word is adapted with a long final vowel (rather than a short one). The independent variable of interest is the status of the preceding vowel: epenthetic, corresponds to unstressed in English, or corresponds to stressed in English:[13]

(21) Logistic regression model for final V length. Dependent variable is whether final V is adapted as long (vs. short).

| | Coeff. estimate | Std. Error | z value | p | |
|---|---|---|---|---|---|
| (Intercept) | -0.653 | 0.317 | -2.06 | 0.039 | |
| prec. V status | | | | | |
| = epenthetic | 1.712 | 0.488 | 3.51 | < 0.001 | * |
| = corr. to unstressed | *reference level* | | | | |
| = corr. to stressed | -3.469 | 1.211 | -2.86 | 0.004 | * |

The model shows that, compared to the baseline case where the preceding vowel corresponds to an unstressed English vowel, a long final vowel is significantly more likely when the preceding vowel is epenthetic (positive coefficient), and significantly less likely when the preceding vowel corresponds to a stressed English vowel (negative coefficient).

That is, final vowel length is indeed favored if it avoids stress on a preceding vowel that corresponds to an English unstressed vowel, and even more if it avoids stress on a preceding epenthetic vowel.

### 5.3 Discussion

In this case, Tongan speakers are making a three-way distinction among English CV́C/CV̀C, CVC, and CC, even though none of those distinctions are contrastive in Tongan.

What is more surprising is that the length effects are seen not on the vowel that itself is stressed, unstressed, or absent in English, but on the subsequent vowel. It’s not just that stressed vowels are more likely than unstressed to be perceived as long, or that actual vowels are more likely to be perceived as long vowels than C-C transitions are. Rather, a subsequent vowel is being treated as long if doing so prevents stress on a vowel whose English correspondent is unstressed or nonexistent.

## 6 Analysis

We now have three instances where Tongan loan adaptations are sensitive to the English CC vs. CVC contrast, and one instance (vowel length) where they are indirectly sensitive to the English stressed/unstressed contrast, even though neither difference is contrastive in Tongan. What does this mean for the loan-adaptation process, and for the grammar and underlying representations once these words are adapted? We begin with vowel deletion, because the analyses of secondary stress and vowel length will depend on it.

[13] As shown in the supplemental materials, we began from a model that also includes the independent variables speaker and final vowel quality ([oʊ] or [i] vs. schwa), with maximal interactions. Using the function step(), we determined that the best model, shown in (21), was one with only preceding vowel status as a factor.

The analysis will treat loan adaptation as a series of mappings, following Boersma and Hamann (2009): from an auditory form to a perceived phonological surface form to a lexical representation during listening, and back to a surface form and an articulatory form during speaking. Each step of the mapping will be governed by a Maximum Entropy constraint grammar (Goldwater & Johnson 2003). We will adopt a representational account: words with different behavior will have different underlying representations (e.g., Inkelas, Orgun & Zoll 1996), rather than bearing different diacritics (e.g., Pater 2006).

### 6.1 Vowel deletion

We saw in section 4 that unstressed *i* is more likely to be deleted when epenthetic with respect to English. For example, we observed undeleted *i* in [palès**i**téni] 'pres**i**dent' and [mèkes**i**kóo] 'Mex**i**co', but deleted *i* in some tokens of [kalìstála] 'crys_tal' and [mìnistáa] 'minis_ter'. Overall, non-epenthetic *i* deleted 14% of the time, and epenthetic *i* 46% of the time.

We pursue an analysis with an underlying difference between full vowels and weak vowels, as in *pal***s*****i***$_{\mathit{full}}$*teni*$_{\mathit{weak}}$ vs. *kalis***i***$_{\mathit{weak}}$*tala*, including analysis of how speakers end up with these underlying forms.

We adopt Boersma and Hamann's (2009) model of loan adaptation, where the grammar regulates a series of mappings, as listed in (22). An English utterance like [kɹístəl] first must be mapped by the listener to a perceived surface form, which is the same type of object as a produced surface form, including stress and foot structure: e.g., /ka$_{weak}$(lì$_{full}$si$_{weak}$)(tá$_{full}$la$_{weak}$)/. (In Boersma and Hamann's notation, slashes surround phonological surface forms, and vertical bars surround underlying forms.) The perception grammar governing this mapping is the same one used for native-language perception. To arrive at an underlying form—whether for a known word or, in the case of initial adaptation, a new word whose underlying form must be constructed—a recognition grammar is used, again the same one as for native words. When the listener wishes to say the word, a production grammar maps the underlying form to a produced surface form, and a phonetic implementation grammar maps this surface form into articulation, perhaps realising /i$_{weak}$/ as overlapped or reduced to the point of deletion, as in the example in (22).

(22) Mappings in the Boersma-Hamann model, with examples for 'crystal'

- perception: [auditory form] to perceived /surface form/
  [kɹístəl] → /ka$_{weak}$(lì$_{full}$si$_{weak}$)(tá$_{full}$la$_{weak}$)/
- recognition: perceived /surface form/ to |underlying form|
  /ka$_{weak}$(lì$_{full}$si$_{weak}$)(tá$_{full}$la$_{weak}$)/ → |ka$_{weak}$li$_{full}$si$_{weak}$ta$_{full}$la$_{weak}$|
- phonological production: |underlying form| to produced /surface form/
  |ka$_{weak}$li$_{full}$si$_{weak}$ta$_{full}$la$_{weak}$| → /ka$_{weak}$(lì$_{full}$si$_{weak}$)(tá$_{full}$la$_{weak}$)/
- phonetic implementation: /surface form/ to [articulation]
  /ka$_{weak}$(lì$_{full}$si$_{weak}$)(tá$_{full}$la$_{weak}$)/ → [kalìstála]

Nowhere in this model is there any special mechanism for loans, although some constraints may be more relevant for loans. We combine all of these mappings into a single grammar. The grammar includes markedness constraints that penalise structures in surface forms—applicable to both perception and production—and constraints regulating mappings between each adjacent pair of levels.

Weak vowels can occur in both underlying and surface forms. The full/weak distinction can be thought of as a feature, like voiced/voiceless. Unlike the excrescent or intrusive vowels investigated by Hall (2003, 2006), which originate as gaps between consonant constrictions, Tongan weak vowels do count for footing and stress. (If they didn't, then feet like (té$_{\text{full}}$ni$_{\text{weak}}$) would be problematic—the /n/ would be a coda consonant. An analysis along these lines is possible, but would be different from ours, where NoCoda is undominated.) The idea, however, is similar: there is a perceptual continuum from zero to a full vowel, and the zero in an illegal consonant sequence will tend to get perceived as some kind of vowel (Hall 2006: 9; Davidson 2007; Davidson & Shaw 2012; Broselow 2015).

In our analysis, the difference between epenthetic and non-epenthetic vowels arises in the perceptual mapping. Tongan listeners hear instances of consonants not followed by a full vowel—CC, CV̥C, C#, and CV̥#—in native words, loan words, and English speech. As shown in (23) (where full/weak status is shown only for the key vowel) under a simplified scenario with no variation, high-ranking NoCoda and *ComplexOnset (Prince & Smolensky 1993) force these consonant transitions, consonant releases, and devoiced vowels to be perceived as vowels (similar to Boersma and Hamann's (2009) analysis of epenthesis in Korean loans). Under the ranking shown here, perceiving such sounds as weak vowels is preferable to perceiving them as full vowels.

(23) Perception of consonant releases and transitions as weak vowels (contextual mapping constraints that would determine quality of perceived weak vowel are omitted)

| | [pɹɛ́zədɛ̀nt] | /NoCoda/ | */Complex/ | *[V$_{\text{full}}$] → /V$_{\text{weak}}$/ | *[release] → /V$_{\text{full}}$/ | *[release] →/V$_{\text{weak}}$/ | */V$_{\text{weak}}$/ |
|---|---|---|---|---|---|---|---|
| *a* | /(plè.**si$_{\text{full}}$**)(tén)/ | *(!) | *(!) | | | | |
| *b* | ☞ /pa(lè.**si$_{\text{full}}$**)(té.ni)/ | | | | | | |
| *c* | /pa(lè.**si$_{\text{weak}}$**)(té.ni)/ | | | *! | | | * |
| | [kɹístəl] | /NoCoda/ | */Complex/ | *[ V$_{\text{full}}$ ] → /V$_{\text{weak}}$/ | *[release] → /V$_{\text{full}}$/ | *[release] →/V$_{\text{weak}}$/ | */V$_{\text{weak}}$/ |
| *e* | /(klìs)(tál)/ | *(!)* | *(!) | | | | |
| *f* | /ka(lì.**si$_{\text{full}}$**)(tá.la)/ | | | | *! | | |
| *f* | ☞ /ka(lì.**si$_{\text{weak}}$**)(ta.la)/ | | | | | * | * |

- /NoCoda/: in a surface form, a syllable must not end with a consonant
- /*Complex/: in a surface form, a syllable must not begin with more than one consonant
- *[V$_{\text{full}}$] → /V$_{\text{weak}}$/: A full vowel in the auditory form should not be perceived as a weak vowel
- *[release] → /V$_{\text{full}}$/: A consonant release or transition in the auditory form should not be perceived as a full vowel
- *[release] → /V$_{\text{weak}}$/: A consonant release or transition in the auditory form should not be perceived as a weak vowel
- */V$_{\text{weak}}$/: Surface forms should not contain weak vowels

The term ‘consonant release or transition’ covers a lot of ground, from a stop’s release burst, as in [pɹ], to a sudden decrease in amplitude, as in [st] or [l#]. In more detailed model, constraints would distinguish these cases. However, we did not find systematic effects of consonant sequence,[14] so our analysis does not break down release/transition types.

When there is a devoiced vowel, the weak vowel’s perceived quality is straightforward; when there is no vowel, as in ‘crystal’, the weak vowel’s quality will be whatever is perceived as closest to the consonant transition (see Uffmann 2007 on trends in epenthetic loan vowels).

Assuming that the mapping from perceived surface form to lexical entry is faithful, we then have underlying |palesi$_{full}$teni| and |kalisi$_{weak}$tala|. In production, constraints against weakening underlying full vowels and vice versa maintain the difference:

(24) Faithful production of full and weak vowels

| | \|pales$\mathbf{i}_{\mathbf{full}}$teni\| | *\|$V_{full}$\| → /$V_{weak}$/ | *\|$V_{weak}$\| → /$V_{full}$/ | */$V_{weak}$/ |
|---|---|---|---|---|
| *a* | ☞ /pa(lè.s$\mathbf{i}_{\mathbf{full}}$)(té.ni)/ | | | |
| *b* | /pa(lè.s$\mathbf{i}_{\mathbf{weak}}$)(té.ni)/ | *! | | * |
| | \|kalis$\mathbf{i}_{\mathbf{weak}}$tala\| | *\|$V_{full}$\| → /$V_{weak}$/ | *\|$V_{weak}$\| → /$V_{full}$/ | */$V_{weak}$/ |
| *c* | /ka(lì.s$\mathbf{i}_{\mathbf{full}}$)(tá.la)/ | | *! | |
| *d* | ☞ /ka(lì.s$\mathbf{i}_{\mathbf{weak}}$)(tá.la)/ | | | * |

- *|$V_{full}$| → /$V_{weak}$/: an underlying full vowel should not correspond to a weak surface vowel
- *|$V_{weak}$| → /$V_{full}$/: an underlying weak vowel should not correspond to a surface full vowel

Finally comes phonetic implementation, where weak vowels are pronounced with gestures that are reduced in magnitude and/or duration, leading to devoicing, overlap, and deletion. Again abstracting away from variation, we assume that the /i/ of ‘president’ is realised as a full [i], and the /i$_{weak}$/ of ‘crystal’ is deleted:

(25) Phonetic realization of full and weak medial /i/, ignoring other vowels in word

| | /pa(lè.s$\mathbf{i}_{\mathbf{full}}$)(té.ni)/ | */$V_{full}$\|/→ [Ø] | */$V_{weak}$/ → [V] |
|---|---|---|---|
| *a* | ☞ [palès**i**téni] | | |
| *b* | [palèsténi] | *! | |
| | /ka(lì.s$\mathbf{i}_{\mathbf{weak}}$)(tá.la)/ | */$V_{full}$\|/→ [Ø] | */$V_{weak}$/ → [V] |
| *c* | [kalìs**i**tála] | | *! |
| *d* | ☞ [kalìstála] | | |

[14] The supplemental R markdown file gives a breakdown of deletion rate by English consonant transition type (for the epenthetic vowel). There are only three categories with a good amount of data, fricative__stop, nasal__stop, and sonorant__fricative. Sonorant__fricative has a lower deletion rate than the other two transition types.

- */$V_{full}$/→ [Ø]: a surface full vowel should not be realised as null
- */$V_{weak}$/ → [V]: a surface weak vowel should not be realised as a full vowel

Weak vowels thus contrast with full vowels in that both can appear in underlying forms, and active faithfulness constraints discriminate between them. But the contrast is marginal in that the faithfulness constraints are too weak to produce stably different pronunciations, and instead only affect variation rates.

To model variation, we fitted a Maximum Entropy constraint grammar (Goldwater & Johnson 2003, Hayes & Wilson 2008, Jäger 2007), shown in (26), using the MaxEnt Grammar Tool (Hayes, Wilson & George 2009)[15]. Each constraint has a weight, and a candidate's probability is proportional to the weighted sum of its constraint violations, exponentiated. For example, for the two candidates in the last tableau of (26), the exponentiated weighted constraint violation sums are 0.0025 and 0.0033; *k*'s proportion of the total, 0.0025/0.0058 = 0.43, is *k*'s probability. The last two columns show the fitted model's predicted probability, and the rate observed in our data across all three speakers. The grammar was fitted to the full data—vowel deletion, secondary stress, and vowel length—but only the relevant constraints are shown here. The full grammar is provided in the supplementary materials. (The constraints /NoCoda/ and */ComplexOnset/ are omitted, on the assumption that they have such high weights as to exclude any candidates that violate them.)

The tableaux in (26) illustrate the mappings from auditory form to surface form (a-d) to underlying form (e-h)), and back to produced surface form (i-l). An additional markedness constraint, */Unstressed$V_{full}$/, gives even underlying full vowels a chance of being produced as weak when unstressed. Because 'president' and 'crystal' are only representatives of groups of words, we ignore the status of vowels in the word besides the target, and ignore constraint violations that those other vowels incur.

[15] We used default values for every constraint: $\mu$=0 and $\sigma^2$=100,000 (maximal fitting to data).

(26) MaxEnt model of vowel deletion (represented as $V_{weak}$ in surface forms)

| | | *[release] → /$V_{full}$/ | *[release] → /$V_{weak}$/ | *[ V] → /$V_{weak}$/ | */$V_{full}$/ → \|$V_{weak}$\| | */$V_{weak}$/ → \|$V_{full}$\| | *\|$V_{full}$\| → /$V_{weak}$/ | *\|$V_{weak}$\| → /$V_{full}$/ | */UNSTRESSED$V_{full}$/ | */$V_{weak}$/ | wtd sum | $e^{-wtd\_sum}$ | prob. | trained rate |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *weights* | 5.8 | 0.0 | 12.3 | 14.5 | 14.1 | 0.5 | 1.4 | 4.6 | 6.0 | | | | |
| | [pɹɛ́zədɛ̀nt] | | | | | | | | | | | | | |
| *a* | /pa(lè.s$\mathbf{i_{full}}$)(té.ni)/ | | | | | | | | * | | 4.6 | 0.0101 | 1.00 | 1.00 |
| *b* | /pa(lè.s$\mathbf{i_{weak}}$)(té.ni)/ | | | * | | | | | | * | 18.3 | 0.0000 | 0.00 | 0.00 |
| | /pa(lè.s**i**)(té.ni)/ | | | | | | | | | | | | | |
| *c* | \|pales$\mathbf{i_{full}}$teni\| | | | | | | | | | | 0.0 | 1.0000 | 1.00 | 1.00 |
| *d* | \|pales$\mathbf{i_{weak}}$teni\| | | | | * | | | | | | 14.5 | 0.0000 | 0.00 | 0.00 |
| | \|pales$\mathbf{i_{full}}$teni\| | | | | | | | | | | | | | |
| *e* | /pa(lè.s$\mathbf{i_{full}}$)(té.ni)/ | | | | | | | | * | | 4.6 | 0.0101 | 0.86 | **0.84** |
| *f* | /pa(lè.s$\mathbf{i_{weak}}$)(té.ni)/ | | | | | | * | | | * | 6.4 | 0.0016 | 0.14 | **0.16** |
| | [kɹístəl] | | | | | | | | | | | | | |
| *g* | /ka(lì.s$\mathbf{i_{full}}$)(tá.la)/ | * | | | | | | | * | | 10.4 | 0.0000 | 0.01 | 0.00 |
| *h* | /ka(lì.s$\mathbf{i_{weak}}$)(tá.la)/ | | * | | | | | | | * | 6.0 | 0.0025 | 0.99 | 1.00 |
| | /ka(lì.s$\mathbf{i_{weak}}$)(tá.la)/ | | | | | | | | | | | | | |
| *i* | \|kalis$\mathbf{i_{full}}$tala\| | | | | | * | | | | | 14.1 | 0.0000 | 0.00 | 0.00 |
| *j* | \|kalis$\mathbf{i_{weak}}$tala\| | | | | | | | | | | 0.0 | 1.0000 | 1.00 | 1.00 |
| | \|kalis$\mathbf{i_{weak}}$tala\| | | | | | | | | | | | | | |
| *k* | /ka(lì.s$\mathbf{i_{full}}$)(tá.la)/ | | | | | | | * | * | | 6.0 | 0.0025 | 0.50 | **0.54** |
| *l* | /ka(lì.s$\mathbf{i_{weak}}$)(tá.la)/ | | | | | | | | | * | 6.0 | 0.0025 | 0.50 | **0.46** |

- */UNSTRESSED$V_{full}$/: in a surface form, a full (non-weak) vowel should not be unstressed

The challenge to fitting a Boersma-Hamann model is that variation could occur at every level. That is, a produced surface form /ka(lì.si$_{weak}$)(tá.la)/ might come from underlying |kalisi$_{weak}$tala| or |kalisi$_{full}$tala|; and each of those might have been lexicalised from perceived /ka(lì.si$_{weak}$)(tá.la)/ or /ka(lì.si$_{full}$)(tá.la)/. All we can observe is the rate of vowel deletion/weakening, not the intermediate-stage rates of variation. Boersma (2011) calls this the problem of 'whole-language simulation'.[16] We simplify by assuming that absent vowels in English are always perceived and lexicalised as weak vowels, and the variation resides entirely in production. (Thus, the learner is trained on rates of 1.00 and 0.00 for the first two steps of each mapping.) It is also a simplification to assume that weak vowels are deleted only at the final stage of phonetic implementation. All this simplification prevents us from achieving a close numerical fit, but recall that exact rates varied among the three speakers too (see (14)). The model succeeds in capturing the difference between *president*, with a lower deletion rate, and *crystal*, with a higher rate.

### 6.1.1 More on the full/weak distinction

Before moving on to consider other analyses of vowel deletion, we take up questions related to our proposal of a full/weak distinction in vowels.

Nothing in this grammar is specific to loanwords—the same mechanisms would apply to native words. What is surprising about this analysis is that it requires listeners to have perceived a difference between a consonant release and a vowel, even though there was no such contrast in Tongan before English loans entered the language. If we had set out to analyze devoicing and deletion in native words only, we might have located the process in the mapping from phonological forms, like /(pòŋi)(póŋi)/ 'morning', to articulatory forms such as [pòŋpóŋi]—there would be no need for weak vowels, nor for constraints that regulate when they are perceived. One reason not to locate devoicing and deletion in phonetic implementation is Feldman's (1978) observation that weakening is more common before a morpheme boundary—the sort of effect that should take place at the phonological level, not in articulation.[17] If we do locate vowel weakening in the surface form, as in /(pòŋi$_{\text{weak}}$)(póŋi)/, there is the advantage that weak vowels can now be used directly as a cue to stress. Stressed Tongan vowels appear never to devoice or delete; in our analysis this is encoded by a markedness constraint */V́$_{\text{weak}}$/ (see section 6.2). In running speech, */V́$_{\text{weak}}$/ contributes to ruling out an incorrect word segmentation such as */po(ŋì$_{\text{weak}}$po)#ŋi.../.[18]

Our analysis predicts that native words, like loans, can also have weak vowels in their underlying forms. If a listener hears *poŋipoŋi* 'morning' for the first time, and it is pronounced [poŋpoŋi], she could conclude that the underlying form is |poŋV$_{\text{weak}}$poŋi|. But if she subsequently encounters tokens [poŋipoŋ] and [poŋipoŋi], she might revise that underlying form. A weak vowel in an underlying form is thus stable only to the extent that the word's pronunciation is stable. There will be variation across speakers, and maybe uncertainty about some words. As a result, there may be an underlying contrast between weak and full vowels—in the sense that some vowels in some words idiosyncratically undergo more vowel deletion/devoicing, because they are underlyingly

[16] What is needed is a Hidden Markov Model whose transition matrix is regulated by a MaxEnt classifier, as in McCallum, Freitag & Pereira (2000) but with some modifications necessary. See Zuraw & Huynh (submitted) for an implementation, including fit to the Tongan data.

[17] In a modular framework like that argued for by Bermudez-Otero (2012); on the other side, Kawahara 2011 reviews research supporting a phonetics-morphology interface.

[18] If vowel weakening happens in articulation, and there are no weak vowels in phonological surface forms, then we would need two constraints, *[release] → /V́/ to rule out perceiving a release as a stressed vowel, and */V́/ → [Ø or V̥] to rule out producing a stressed vowel as devoiced or deleted.

weak for most speakers—but the contrast is unstable, and difficult for both learners and linguists to observe.

Are loans any different? It is unknown whether the degree of within-word consistency is similar for loans and native words. Finding out would require a large-scale production study. We could expect greater stability in loans, because bilingual speakers encounter the English pronunciation from time to time, and that portion of their input is stable. For example, there may be some uncertainty as to whether the underlying form of 'president' is /pa$_{\text{weak}}$les**i**$_{\textbf{weak}}$teni$_{\text{weak}}$/ or /pa$_{\text{weak}}$les**i**teni$_{\text{weak}}$/, but hearing English [pɹɛ́zədɛ̀nt] reinforces /pa$_{\text{weak}}$les**i**teni$_{\text{weak}}$/.

Our proposal predicts that a weak/full distinction is available in all languages—that is, grammars consider candidates with weak vowels, even if they are not optimal. Are there other languages besides Tongan where weak and full vowels plausibly occur in actual, optimal surface forms? In several languages, some or all schwas have been analyzed as non-moraic, which is akin to our category of 'weak' (e.g., German: Féry 1991; Kabardian: Peterson 2007; Salish languages: (Kinkade 1998) A possible case that does not have to do with moraicity is Japanese, as discussed in section 1. To recap, Kubozono (2006: 1147) shows that epenthetic loan vowels are treated differently, repelling pitch accent: English /pléɪ/ 'play' is adapted as /p**u**rée/, but English /pʊ́ɾɪŋ/ 'pudding' as /p**ú**riN/. Because pitch accent is contrastive in Japanese, the lexical entries for these words don't need to encode the full/weak distinction—they should simply encode the pitch accent: |purée|, |púriN|, etc. But a plausible mechanism for getting to those lexical entries is that the English words are perceived as the phonological surface forms /pu$_{\text{weak}}$rée/, /pú$_{\text{full}}$riN/. In native Japanese words too, perceiving vowels as weak or full could contribute to accurately perceiving a vowel's phonological accentedness and other properties.

Plausible candidates for a weak/strong distinction in *lexical entries* are any languages with lexical exceptions to vowel reduction processes and the like. For example, the vowel reduction that normally applies to Eastern Catalan unstressed vowels has lexical exceptions (Mascaró 2002, as cited by Cabré 2009, who gives a different analysis based on lexical strata). An exception like [dém**o**] 'demonstration' (expected [dém**u**]) could have the lexical entry |démo$_{\text{full}}$|, with a constraint prohibiting reduction of full vowels. Féry (1991: 66, fn. 2) states that whether a German schwa is moraic or not probably must be marked in its lexical entry; a possible re-analysis would be that there is an underlying weak/full distinction, with moras allocated in the phonological surface form accordingly.

Are two degrees of vowel strength (full and weak) enough? It would be difficult to get evidence for finer distinctions. The evidence needed in Tongan would be at least three different words, used by the same speaker with different rates of vowel deletion. For example, hypothetical *san**i**palo* with 25% deletion, *man**i**pako* with 50% and *tan**i**pano* with 75%. The words would have to be alike in every relevant respect: morphological structure, stress pattern, consonant environment, and relation to loan source if any. They would have to be sampled from similar distributions of register, sentence position, and focus. Just as such data would be difficult for the linguist to obtain, they would be difficult for the learner to obtain. We conjecture that even if a speaker had constructed a lexicon with more than two degrees of vowel strength, it would be difficult to pass that lexicon on faithfully to a learner. On the other hand (going beyond the scope of this paper), if we abandoned symbolic lexical entries altogether and adopted exemplar representations (Johnson 1997; Pierrehumbert 2001), each word would have its own cloud of exemplars, with various degrees of reduction. There would then be a vast range of effective vowel strengths, because of the vast range of possible exemplar clouds.

### 6.1.2 Other analyses of vowel deletion

The other obvious analysis of vowel deletion is that English (and Tongan) consonant clusters are simply perceived as such. That is, NOCODA and *COMPLEXONSET are not ranked high enough to force perception of a vowel, and even if weak vowels are possible phonological symbols, they will not arise in Tongan, because *[release] → /V/ prevents a consonant release from being perceived as any vowel, weak or full:

(27) Perceiving null as null

| | [pɹɛ́zədɛ̀nt] | *[release] → /V/ | /NOCODA/ |
|---|---|---|---|
| *a* | ☞ /pa(lè.s**i**)(té.ni)/ | | |
| *b* | /pa(lès)(té.ni)/ | | *! |
| | [kɹístəl] | *[release] → /V/ | /NOCODA/ |
| *c* | /ka(lì.s**i**)(tá.la)/ | *! | |
| *d* | ☞ /ka(lìs)(tá.la)/ | | * |

These surface forms would be lexicalised faithfully as |palesiteni| and |kalistala|, and produced faithfully as /pa(lè.si)(té.ni)/ and /ka(lìs)(tá.la)/. Variation would arise in phonetic realization (or the underlying-to-surface mapping), with medial unstressed /i/ sometimes getting reduced in articulation—including to the point of deletion—and with surface consonant clusters like /st/ sometimes being realised with a vocalic gap or excrescent vowel (Hall 2003; Hall 2006) so robust as to resemble a true vowel.

A more radically different analysis appeals to output-output-correspondence (Kenstowicz 1996, Benua 1997, Crosswhite 1998, Burzio 1999, Steriade 2000). 'President' and 'crystal' would have similar underlying forms, but bilinguals—like our consultants—would consult the English pronunciation directly, as shown in (28). The constraints labeled 'IO' regulate the mapping between the input and the output, and those labeled 'OO' regulate the mapping between the English pronunciation and the output. MAX-V penalises deleting a vowel, and DEP-V penalises inserting one (McCarthy & Prince 1995); subscript numerals show the assumed correspondence relations. An approach like this has been employed for loan adaptation by Kenstowicz (2005), and Smith (2006a), who both point out that speakers may draw on information from spelling to help evaluate OO constraints (here, the fact that there is a vowel letter between the *s* and *d* of *president*, but not between the *s* and *t* of *crystal*).[19]

[19] Daland, Oh & Kim (2015) devised an ingenious way to compare the contributions of pronunciation and spelling when they conflict, but their approach is not applicable here, because English spelling is generally a good guide to whether a word-internal vowel is present or not—there is no disagreement between spelling and pronunciation that would allow us to compare their contributions.

(28) Output-output correspondence in production (no need for Boersma-Hamann model)

| | pales$_1$i$_2$t$_3$eni, cf. [pɹέz$_1$**ə$_2$**d$_3$ɛ̀nt] | MAX-V-OO | DEP-V-OO | MAX-V-IO | DEP-V-IO |
|---|---|---|---|---|---|
| *a* | ☞ pa(lè.s$_1$**i$_2$**)(t$_3$é.ni) | | | | |
| *b* | pa(lès$_1$)(t$_3$é.ni) | *! | | * | |
| | kalis$_1$i$_2$t$_3$ala, cf. [kɹís$_1$t$_3$əl] | MAX-V-OO | DEP-V-OO | MAX-V-IO | DEP-V-IO |
| *c* | ka(lì.s$_1$**i$_2$**)(t$_3$á.la) | | *! | | |
| *d* | ☞ ka(lìs$_1$)(t$_3$á.la) | | | * | |

The OO analysis makes a prediction that our data do not support: speakers should differ in the strength of the OO correspondence effect. Some speakers—through greater English proficiency, or greater tendency to retrieve the English form during speech—should show more disparity between epenthetic and non-epenthetic vowels' deletion rates, and some should show less. (Both the OO analysis and our analysis could *capture* such between-speaker differences by simply having different constraint weights across speakers, but the OO analysis particularly *predicts* such differences.) We did not observe this. Instead, the three speakers differed in their overall rate of deletion, capturable through differing weights of */Vfull/ in our analysis, or something like *V in the OO analysis. Perhaps a larger group of speakers, with a wider range of English proficiencies (our consultants had lived in the U.S. for years), would show the differences predicted by the OO analysis. But if the cross-speaker pattern in our data proves representative, it is an argument against the OO analysis.

### 6.2 Secondary stress

In section 3, we saw that if the first vowel in a loan is epenthetic, it is less likely to bear secondary stress. When neither of the first two vowels in a loan (that meets the shape criteria in 3.2) is epenthetic, secondary stress falls on the first vowel 68% of the time, as in [(tèmo)ka(láti)] 'demócracy'. We therefore take initial stress as the default, illustrated in (29). ALIGN constraints (McCarthy & Prince 1993) govern foot placement: ALIGN(PWord,R; Foot,R) requires every prosodic word to end with a foot, and ALIGN(PWord,L; Foot,L) requires every word to begin with a foot.[20] Countervailing ALIGN(Foot,R; Pword,R) requires all feet to align towards the right, but is less powerful (variation is modeled below).

(29) Default stress pattern: initial dactyl

| | \|temokalati\| | ALIGN (PWord,R; Foot,R) | ALIGN (PWord,L; Foot,L) | ALIGN (Foot,R; PWord,R) |
|---|---|---|---|---|
| *a* | ☞ /(tèmo)ka(láti)/ | | | *** |
| *b* | /te(mòka)(láti)/ | | *! | ** |
| *c* | /(tèmo)(kála)ti/ | *! | | ***,* |

[20] ALIGN(Foot,L; PWord,L) would also do the job, by requiring every foot to be close to the beginning of the prosodic word. This constraint would also predict that if there were a third foot it would align to the left: (tata)(tata)ta(tata) instead of (tata)ta(tata)(tata). But we lack sufficient words of that shape or a reliable way of detecting the position of the medial stress.

Words like [falàkiséni] ‘fraction’, whose first vowel is epenthetic, depart from this pattern, with only 26% having initial secondary stress. In the analysis developed above in section 6.1, English *fraction* [fɹǽkʃən] should be perceived as /f$\mathbf{a}_{\mathbf{weak}}$(làki)(séni)/, with a weak first vowel, whereas *democracy* [dɨmɒ́krəsi] should be perceived as /(t$\mathbf{è}_{\mathbf{full}}$mo)ka(lasi)/.

The tableau in (30) illustrates the mapping from auditory to perceptual form, simplified to categorical outcomes. We add one new markedness constraint, */V́$_{weak}$/, forbidding stressed, weak vowels in surface forms. This is reminiscent of Kenstowicz's (2007) two-level constraint *´v, which penalises stressing an epenthetic vowel. The advantage of a markedness constraint like */V́$_{weak}$/ is that it can regulate both perception and production. In production, it prevents a stressed vowel from undergoing reduction (and therefore is needed in native Tongan words too). In perception, it ensures that a reduced vowel is not perceived as stressed, which is also a reasonable strategy for native Tongan words.[21]

(30) Perception of initial CC vs. CVC

| | [fɹǽkʃən] | *[release] → /V$_{full}$/ | */V́$_{weak}$/ | ALIGN (PWord,L; Foot,L) | *[release] →/V$_{weak}$/ | ALIGN (Foot,R; PWord,R) |
|---|---|---|---|---|---|---|
| *a* | ☞ /f$\mathbf{a}_{\mathbf{weak}}$(làki)(séni)/ | | | * | *** | ** |
| *b* | /(f$\mathbf{à}_{\mathbf{weak}}$la)ki(séni)/ | | *! | | *** | *** |
| *c* | /(f$\mathbf{à}_{\mathbf{full}}$la)ki(séni)/ | *! | | | ** | *** |
| | [dɨmɒ́krəsi] | *[release] → /V$_{full}$/ | */V́$_{weak}$/ | ALIGN (PWord,L; Foot,L) | *[release] →/V$_{weak}$/ | ALIGN (Foot,R; PWord,R) |
| *d* | /t$\mathbf{e}_{\mathbf{full}}$(mòka)(lási)/ | | | *! | * | ** |
| *e* | ☞ /(t$\mathbf{è}_{\mathbf{full}}$mo)ka(lási)/ | | | | * | *** |

- */V́$_{weak}$/: weak, stressed vowels in surface forms are penalised

The ranking in (30) allows Tongan speakers to perceive unpredictable placement of secondary-stress feet. This is needless for native monomorphemes in isolation, but a good strategy for comprehension of running speech, where word and morpheme boundaries are not known ahead of time, and stress is a cue to boundaries.

How will this translate into an underlying form in a Boersma-Hamann-style comprehension system? Given what we have already, the simplest answer is that the weak vowels are lexicalised as such: |f$\mathbf{a}_{\mathbf{weak}}$lakiseni| versus |t$\mathbf{e}_{\mathbf{full}}$mokalasi|. In production, */V́$_{weak}$/ and *|V$_{weak}$| → /V$_{full}$/ ensure that secondary stress continues to avoid the weak vowel in ‘fraction’.

[21] It is possible for an epenthetic vowel to end up stressed, as in [ʔiisíti] ‘east’, from English [i:st]. One analysis is to have [i:st] perceived as /(ʔii)(s$\mathbf{í}_{\mathbf{full}}$ti)/, violating *[release] → /V$_{full}$/ in order to satisfy ALIGN(PWord,R; Foot,R) and */V́$_{weak}$/. /(ʔii)(s$\mathbf{í}_{\mathbf{full}}$ti)/ then gets the lexical entry |ʔiisi$_{full}$ti|, just as though the vowel were not epenthetic. Another analysis ranks ALIGN(PWord,R; Foot,R) and *[release] → /V$_{full}$/ above */V́$_{weak}$/, so that even though ‘east’ is lexicalised as |ʔiisi$_{weak}$ti|, it must be produced as /(ʔii)(s$\mathbf{í}_{\mathbf{weak}}$ti)/.

We provide a MaxEnt model of this process. We simplify by assuming that variation happens in production rather than perception. We distinguish three cases: first vowel is epenthetic (‘b_lancmánge’, ‘f_ráction’), second vowel is epenthetic (‘as_trònomer’, ‘kàn_garóo’), or neither is epenthetic (‘demócracy’, ‘mónogràm’). For the sake of space, we show only the highlights of the model. (The full model of all three case studies is in the supplementary materials.) High-weighted *[release] → /$V_{full}$/ ensures that consonant releases are perceived as weak vowels, and high-weighted *[$V_{full}$] → /$V_{weak}$/s ensures that English vowels are perceived as full. At the recognition stage, high-weighted */$V_{full}$/ → |$V_{weak}$| and */$V_{weak}$/ → |$V_{full}$| ensure that these decisions about weakness are lexicalised.

The tableaux in (31) show the variation that then occurs in production. We consider four candidates for each word: peninitial stress or initial stress (on a full vowel) with the other, unstressed vowel weak or full. Under our assumptions about where in the grammar vowel deletion occurs, a full underlying vowel has a 16% chance of being produced weak, and a weak underlying vowel a 46% chance. We divide up the stress rates accordingly: for example, the 84% observed rate of initial stress for *kan_garoo*-type words (second vowel weak: |$\text{ka}_{full}\text{ŋi}_{weak}$kaluu|) becomes a trained rate of 0.46 × 0.84 = 0.39 for /($\text{kà}_{full}\text{ŋi}_{weak}$)ka(lúu)/ and 0.54 × 0.84 = 0.45 for /($\text{kà}_{full}\text{ŋi}_{full}$)ka(lúu)/. To assess the success of the model, we show the combined model probabilities and combined trained rates for the two peninitial-stress candidates and the two initial-stress candidates.

(31) MaxEnt model of secondary stress (weak/full status of vowels other than first two not shown)

| \|ka$_{full}$ŋi$_{weak}$kaluu\| | *\|V$_{full}$\| → /V$_{weak}$/ | *\|V$_{weak}$\| → /V$_{full}$/ | */UNSTRESSEDV$_{full}$/ | */V$_{weak}$/ | ALIGN(PWd, L; Ft, L) | ALIGN(Ft, R; PWd, R) | wtd sum | e$^{-(wt_sum)}$ | prob. | | trained rate | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 0.5 | 1.4 | 4.6 | 6.0 | 1.0 | 0.0 | | | | | | |
| *a* /ka$_{weak}$(ŋì$_{full}$ka)(lúu)/ | * | * | | * | * | ** | 8.8 | 0.0001 | 0.02 | 0.17 | 0.03 | 0.16 |
| *b* /ka$_{full}$(ŋì$_{full}$ka)(lúu)/ | | * | * | | * | ** | 7.0 | 0.0009 | 0.15 | | 0.14 | |
| *c* **/(kà$_{full}$ŋi$_{weak}$)ka(lúu)/** | | | | * | | *** | 6.0 | 0.0025 | 0.41 | 0.83 | 0.39 | 0.84 |
| *d* /(kà$_{full}$ŋi$_{full}$)ka(lúu)/ | | * | * | | | *** | 6.0 | 0.0025 | 0.41 | | 0.45 | |
| \|te$_{full}$mo$_{full}$kalasi\|[22] | | | | | | | | | | | | |
| *e* /te$_{weak}$(mò$_{full}$ka)(lási)/ | * | | | * | * | ** | 7.5 | 0.0006 | 0.04 | 0.27 | 0.06 | 0.34 |
| *f* /te$_{full}$(mò$_{full}$ka)(lási)/ | | | * | | * | ** | 5.6 | 0.0036 | 0.23 | | 0.28 | |
| *g* /(tè$_{full}$mo$_{weak}$)ka(lási)/ | * | | | * | | *** | 6.4 | 0.0016 | 0.10 | 0.73 | 0.11 | 0.66 |
| *h* /(tè$_{full}$mo$_{full}$)ka(lási)/ | | | * | | | *** | 4.6 | 0.0101 | 0.63 | | 0.55 | |
| \|fa$_{weak}$la$_{full}$kiseni\| | | | | | | | | | | | | |
| *i* /fa$_{weak}$(là$_{full}$ki)(séni)/ | | | | * | * | ** | 7.0 | 0.0009 | 0.19 | 0.38 | 0.35 | 0.74 |
| *j* /fa$_{full}$(là$_{full}$ki)(séni)/ | | * | * | | * | ** | 7.0 | 0.0009 | 0.19 | | 0.39 | |
| *k* /(fà$_{full}$la$_{weak}$)ki(séni)/ | * | * | | * | | *** | 7.8 | 0.0004 | 0.08 | 0.62 | 0.05 | 0.26 |
| *l* /(fà$_{full}$la$_{full}$)ki(séni)/ | | * | * | | | *** | 6.0 | 0.0025 | 0.53 | | 0.21 | |

The model manages to capture the three-way distinction: ‘kan_garoo’ has the highest rate of initial secondary stress, followed by ‘democracy’, followed by ‘f_raction’, although the model’s rate of initial stress in ‘f_raction’ is much lower than the training rate.

### 6.2.1 Other analyses of secondary stress

As an alternative to allowing weak vowels (or CC and C#) in underlying representations, we could introduce underlying secondary stress: |falàkiseni|, vs. |kàŋikaluu| or unspecified |kaŋikaluu|. As mentioned in section 3.1, native monomorphemes that are five moras or longer are few if any; this could leave Tongan learners uncertain as to whether secondary stress is predictable or lexical. Constraints on the mappings from acoustic to surface form, and from surface to underlying, would have to ensure that epenthetic first vowels result in underlying forms like |falàkiseni|. While this

[22] There is a constraint to be introduced in section 6.3, *[VRELPROM ] → /UNSTRESSED/, on which words like *monokalame* ‘mónogràm’ and *temokalasi* ‘demócracy’ perform differently. Therefore, we include *monokalame* and *temokalasi* as separate cases in our training-data file, but report here their combined results.

analysis seems plausible, given that we already need weak or zero vowels to handle vowel deletion, we do not pursue it.

Another possibility is to index words to differently-ranked versions of ALIGN(PWord,L; Foot, L) (Pater 2000). *Kaŋikaluu* would be indexed to a high-ranking version, and *falakiseni* would be indexed to a low-ranking version. This approach would require a mechanism to determine which version of the constraint a new word gets indexed to (see Pater 2009). Sensitivity to epenthetic-hood would reside in this mechanism rather than in the constraints of the grammar. Again, given that we already need weak or zero vowels to handle vowel deletion, we do not pursue this analysis. (An indexed-constraint approach to vowel deletion is unattractive because indexation would have to be to individual vowels rather than to whole morphemes, contradicting one of the assumptions of the theory of constraint indexation as it has been developed by Pater and colleagues; though cf. Round 2017)

It is also possible to pursue an output-output correspondence analysis. This would be less straightforward than the output-output analysis of vowel deletion sketched in 6.1.2, because we don't need faithfulness to the English stress, but rather a penalty for stressing an epenthetic vowel (similar to Kenstowicz's 2007 two-level constraint *´v, which forbids stressing an epenthetic vowel, but assessed relative to the English pronunciation rather than the input):

(32) Output-output correspondence for secondary stress

| | f$_1$a$_2$l$_3$akiseni, cf. [f$_1$ɹ$_3$ǽkʃən] | *STRESS-EPENTHETIC-OO | ALIGN (PWord,L; Ft, L) | ALIGN (Foot,R; PWord,R) |
|---|---|---|---|---|
| *a* | (f$_1$à$_2$l$_3$a)ki(séni) | *! | | *** |
| *b* | ☞ f$_1$a$_2$(l$_3$àki)(séni) | | * | ** |
| | temokalasi, cf. [dɪˈmɒkrəsi] | *STRESS-EPENTHETIC-OO | ALIGN (PWord,L; Ft, L) | ALIGN (Foot,R; PWord,R) |
| *c* | ☞ (tèmo)ka(lási) | | | *** |
| *d* | te(mòka)(lási) | | *! | ** |
| | kaŋikaluu, cf. [kaŋgərúː] | *STRESS-EPENTHETIC-OO | ALIGN (PWord,L; Ft, L) | ALIGN (Foot,R; PWord,R) |
| *e* | ☞ (kàŋi)ka(lúu) | | | *** |
| *f* | ka(ŋìka)(lúu) | *! | * | ** |

- *STRESS-EPENTHETIC-OO: an output vowel that has no correspondent in the English pronunciation should not be stressed

Another possibility is that a word like *fraction* is perceived and lexicalised as a pseudocompound /fa(láki)-(séni)/, with the morphological structure of a compound even though neither 'stem' is a known morpheme in its own right. Consequently, each pseudostem ends in a foot. Pseudocompound analyses have been proposed for loans that would otherwise have illegal stress: Välimaa-Blum (1986) and references therein for Finnish; Hayes (1995) and Árnason (1996) for Icelandic; Munro & Riggle (2004) for Pima; Martin (2005) for Malagasy. In these other languages, unlike in Tongan, there was additional evidence for the pseudocompounding analysis (from vowel harmony or reduplication).

### 6.3 Vowel length

We saw in Section 5 that a final vowel was always adapted as short if this would place stress on a the correspondent of a stressed English vowel (sèpit**é**m**a** ‘Sept**é**mb**er**’), less often if it stress the correspondent of an unstressed English vowel (tàiam**í**t**a** ‘diám**e**t**er**’, kàsit**o**m**áa** ‘cúst**o**m**er**’), and even less often if it would place stress on an epenthetic vowel (*mìnis**í**t**a**, rather minis**itáa** ‘mínis_t**er**’).

The analysis developed so far extends easily to these data. As shown in (33) (ignoring variation for now), the transition from [s] to [t] is perceived as a weak vowel, and the markedness constraint */V́$_{weak}$/, introduced in 6.2, prevents stress from falling on it. The remaining options are a non-binary foot (d), which is not allowed in Tongan, or perceiving the final vowel as long (f).

(33) Auditory-to-surface perception mapping for ‘minister’—footing of beginning of word is ignored

| | [mínɪstə] | *[release] → /V$_{full}$/ | /FOOTBIN/ | */V́$_{weak}$/ | *[V] → /VV/ |
|---|---|---|---|---|---|
| *a* | /mini(síta)/ | *! | | | |
| *b* | /mini(sí$_{weak}$ta)/ | | | *! | |
| *c* | /minisi(tá)/ | *(!) | *(!) | | |
| *d* | /minisi$_{weak}$(tá)/ | | *! | | |
| *e* | /minisi(táa)/ | *! | | | * |
| *f* | ☞ /minisi$_{weak}$(táa)/ | | | | * |

- /FOOTBIN/: a foot must contain exactly two moras
- *[V] → /VV/: a short vowel in the auditory form must not be perceived as long

This is a suitable perception strategy for native words, too: a devoiced or deleted penult vowel must be unstressed, and therefore must be followed by a long-voweled ultima. The mapping from /minisi(táa)/ to underlying |minisitaa| is then straightforward: Tongan lacks productive vowel-length alternations, so long surface vowels should always be recognised as long underlying vowels.[23]

In order to capture the difference between ‘Septémber’ and ‘cústomer’, we need a constraint sensitive to English stress. We propose one whose sensitivity is indirect: *[VRELPROM ] → /UNSTRESSED/, which requires a vowel in the auditory form that is more prominent than either of its within-word neighbors to be perceived as stressed. This constraint is inspired by Broselow's (2009) FINALVLONG (a final vowel more prominent than the preceding vowel should be recognised as long), but applies more broadly, to vowels anywhere in the word. Like Broselow, we leave open what phonetic properties of an English word are heard as a ‘prominence peak’ (217). English and Tongan stressed vowels have many of the same acoustic properties: higher duration,

[23] Or, in definite forms, as a sequence of a final underlying short vowel and an empty mora from the definite enclitic—see Anderson & Otsuka (2006) and references therein.

energy, and pitch (depending on intonation), and lower F1 (Garellek & White 2015 for Tongan, and references therein for English). Therefore, we can assume that, using native-language cues, Tongan speakers tend to perceive English stressed vowels as prominent.

Obeying this constraint is a reasonable perception strategy for Tongan overall: a vowel that is appreciably more prominent than either of its immediate neighbors within the same word must be stressed—either it is stressed and the neighbor is unstressed, or both are stressed but the vowel in question bears primary stress and its neighbor secondary.

As illustrated schematically (leaving ‘customer’ in a tie) in (34), *[VRELPROM ] → /UNSTRESSED/ requires stressing the penult in ‘September’, because the [έ] that it corresponds to is more prominent than both flanking vowels in the auditory form. In ‘customer’, the constraint is irrelevant—all candidates satisfy it—so there is no need to ensure stress on the penult by having a short final vowel.

(34) Perception of stressed vs. unstressed penult (ignoring earlier foot and status of vowels earlier in the word)

| | [sɛptɛ́mbə] | *[VRELPROM ] → /UNSTRESSED/ | /FOOTBIN/ | *[ $V_{full}$ ] → /$V_{weak}$/ | */$\acute{V}_{weak}$/ |
|---|---|---|---|---|---|
| *a* | ☞ /sepi(téma)/ | | | | |
| *b* | /sepi(té$_{weak}$ma)/ | | | *! | * |
| *c* | /sepite(má)/ | *(!) | *(!) | | |
| *d* | /sepite$_{weak}$(má)/ | *(!) | *(!) | *(!) | |
| *e* | /sepite(máa)/ | *! | | | |
| *f* | /sepite$_{weak}$(máa)/ | *(!) | | *(!) | |
| | [kʌ́stəmə] | *[VRELPROM ] → /UNSTRESSED/ | /FOOTBIN/ | *[ $V_{full}$ ] → /$V_{weak}$/ | */$\acute{V}_{weak}$/ |
| *g* | ☞ /kasi(tóma)/ | | | | |
| *h* | /kasi(tó$_{weak}$ma)/ | | | *! | * |
| *i* | /kasito(má)/ | | *! | | |
| *j* | /kasito$_{weak}$(má)/ | | *(!) | *(!) | |
| *k* | ☞ /kasito(máa)/ | | | | |
| *l* | /kasito$_{weak}$(máa)/ | | | *! | |

- *[VRELPROM ] → /UNSTRESSED/: a vowel that is more prominent than either of its neighbors within the same word must not be perceived as unstressed

Our MaxEnt grammar’s model of variation is shown in (35) (evaluating only the last two vowels of each word). In our analyses above we simplified by confining variation to the production mapping, but in Tongan vowel length is contrastive and not subject to either free variation or systematic alternation, so we have instead put variation in vowel length in perception.

(35) MaxEnt model of final lengthening

| | [mínɪstə] | *[VRelProm] → /Unstressed/ | *[V] → /VV/ | *[ release] → /V$_{full}$/ | *[ release] → /V$_{weak}$/ | *[ V$_{full}$ ] → /V$_{weak}$/ | */UnstressedV$_{full}$/ | */V$_{weak}$/ | */StressedV$_{weak}$/ | FootBin | wtd sum | e$^{-(wt_sum)}$ | prob. | | actual rate |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 0.2 | 2.0 | 5.8 | 0.0 | 12.3 | 4.6 | 6.0 | 10.1 | 14.2 | | | | | |
| *a* | /mini(s**í**t**a**)/ | * | | * | | | * | | | | 10.6 | 0.0000 | 0.07 | 0.07 | 0.25 |
| *b* | /minis**i**(t**á**)/ | | | * | | | * | | | * | 24.6 | 0.0000 | 0.00 | | |
| *c* | /minis**i**$_{weak}$(t**á**a)/ | | * | | * | | | * | | | 8.0 | 0.0003 | 0.92 | 0.93 | 0.75 |
| *d* | /minis**i**(t**á**a)/ | | * | * | | | * | | | | 12.4 | 0.0000 | 0.01 | | |
| | [kʌ́stəmə] | | | | | | | | | | | | | | |
| *e* | /kasi(t**ó**m**a**)/ | | | | | | * | | | | 4.6 | 0.0170 | 0.88 | 0.88 | 0.65 |
| *f* | /kasit**o**(m**á**)/ | | | | | | * | | | * | 18.8 | 0.0000 | 0.00 | | |
| *g* | /kasit**o**$_{weak}$(m**á**a)/ | | * | | | * | | * | | | 20.3 | 0.0000 | 0.00 | 0.12 | 0.35 |
| *h* | /kasit**o**(m**á**a)/ | | * | | | | * | | | | 6.6 | 0.0013 | 0.12 | | |
| | [sɛptɛ́mbə] | | | | | | | | | | | | | | |
| *i* | /sepi(t**é**m**a**)/ | | | | | | * | | | | 4.6 | 0.0101 | 0.99 | 0.90 | 1.00 |
| *j* | /sepit**e**(m**á**)/ | * | | | | | * | | | * | 19.0 | 0.0000 | 0.00 | | |
| *k* | /sepit**e**$_{weak}$(m**á**a)/ | * | * | | | * | | * | | * | 34.7 | 0.0000 | 0.00 | 0.10 | 0.00 |
| *l* | /sepit**e**(m**á**a)/ | * | * | | | | * | | | | 6.8 | 0.0011 | 0.10 | | |

The model captures the three-way distinction of epenthetic, unstressed, and stressed.[24]

### 6.3.1 Other analyses of vowel length

In analyzing a similar final-length phenomenon in Fijian, Kenstowicz (2007) relies on two ingredients. First, an English consonant cluster is lexicalised as such: /minista/, even though the word will be pronounced with an epenthetic vowel because of NoCoda. Second, there is a constraint penalising stress on an epenthetic vowel (*´v), outranking faithfulness to length: thus, length-unfaithful [minisi(táa)] is better than *[mini(síta)]. (Kenstowicz's constraint is a simplified version of Alderete's 1999 Head-Dep, which refers to prosodic head-hood rather than stress. The

[24] Under the constraint set shown, candidate *h* is harmonically bounded by candidate *e*. This is not a problem in MaxEnt, where even harmonically bounded candidates receive some probability; in this case, because the weight of *[V] → /VV/ is small, *h* gets substantial probability.

observation that epenthetic vowels in loans tend to repel stress is due to Broselow 1982; Broselow 2008 provides an overview in OT.)

Broselow (2009) revisits Kenstowicz's analysis of Schütz's Fijian corpus, and points out a puzzle for learnability: when English loans entered the Fijian language, how had Fijian speakers arrived at the production-grammar ranking posited, even though the rankings involved were not needed for native words (which do not undergo epenthesis)?

Broselow adopts a model similar to Boersma & Hamann's (2009), based on Boersma 1998. However, her perception grammar maps auditory [míniɪstə] directly to underlying |ministaa|. This means that rather than perceiving stress, with foot-structure constraints then shaping perceived length, listeners perceive length directly from relative prominence, using FINALVLONG (a final vowel that is more prominent than the preceding vowel should be recognised as long). As in the two-step analysis given above for Tongan, this is a reasonable comprehension strategy for native words: given the many sources of variation in duration, especially word-finally, a good cue to final vowel length—perhaps better than duration alone—is its overall prominence relative to the preceding vowel.

An analysis based on output-output correspondence would be indirect here, as it was for secondary stress. The English final vowels in question are short, so Tongan is not being faithful to their length. We need the constraint *STRESS-EPENTHETIC-OO from section 6.2.1 to ensure a long final vowel in 'minister'. Under strict ranking of these constraints, 'customer' will never have a final long vowel—candidate *d* is harmonically bounded—but in a MaxEnt grammar candidate *d* will still have some probability, and crucially more than *b*.

(36) Output-output correspondence for final length

| | sepitema, cf. [sɛptɛ́mbə] | ID-STRESS-OO | *STRESS-EPENTHETIC-OO | ID-LENGTH-IO |
|---|---|---|---|---|
| *a* | ☞ sepi(téma) | | | |
| *b* | sepite(máa) | *! | | * |
| | kasitoma, cf. [kʌ́stəmə] | | | |
| *c* | ☞ kasi(tóma) | | | |
| *d* | kasito(máa) | | | *! |
| | minisita, cf. [mínɪstə] | | | |
| *e* | mini(síta) | | *! | |
| *f* | ☞ minisi(táa) | | | * |

- ID-STRESS-OO: an output vowel that corresponds to an English stressed vowel must be stressed
- ID-LENGTH-IO: an output vowel must have the same length as its input correspondent

What is unappealing here is that an important, high-ranking constraint, ID-LENGTH-IO, must be weighted lower than *STRESS-EPENTHETIC-OO, a constraint motivated only for loanwords.

## 7 Conclusion

Tongan speakers must be able to perceive the non-native contrast between CVC and CC, well enough to carry it into loan adaptations, though not necessarily in that form. We proposed that the release of a word-final consonant or the transition in a CC sequence is perceived as a weak vowel, and we proposed an analysis in the bidirectional model of Boersma (1998), Boersma & Hamann (2009), with similarities to Kenstowicz's (2007) and Broselow's (2009) analyses of Fijian.

To account for the fact that epenthetic vowels delete more often, we proposed that Tongan allows a category of weak vowels in surface representations. This is not surprising, since native Tongan words also have vowel devoicing, shortening, and deletion. More controversial was the proposal that *underlying* representations could contain weak vowels. In principle native words could have such a contrast too, so that some vowels of some words have an idiosyncratic tendency to undergo vowel weakening more than usual. Variation in pronunciation presumably produces some variation in underlying forms for native and loan words alike, so this contrast is not stable.

Allowing weak vowels in underlying representations also captured why epenthetic vowels repel secondary stress, using a markedness constraint prohibiting stressed, weak surface vowels (*/$\acute{\text{V}}_{\text{weak}}$/). This constraint is needed independently to account for native Tongan phonology, where weakening does not apply to stressed vowels. If Tongan lacked vowel weakening, it would have been more difficult for the secondary-stress effect to arise, because surface forms would not distinguish weak and strong vowels, and even if a word were somehow correctly perceived with aberrant, peninitial secondary stress (/fa(làki)(séni)/ 'fraction'), its lexicalised form (|falakiseni|) would be the same as that of a loan with initial secondary stress.

Our third case study was vowel length: final vowels lengthen to avoid stressing a preceding vowel that was unstressed in English, or, even more strongly, an epenthetic vowel. Taking inspiration from Broselow's analysis of Fijian, we posited a constraint requiring that a vowel be perceived as stressed if it is more prominent than an adjacent vowel—a reasonable part of a stress-perception strategy for native Tongan words. The bidirectional model was crucial to our analysis: the perceptual constraint is active in the mapping from an English auditory form ([mínɪstə]) to a Tongan surface form, complete with feet and stress (/(mìni)si$_{\text{weak}}$(táa)/). The final vowel lengthening is a combined effect of the perception-mapping constraint's stress requirement (*/(mìni)(síta)/) and ordinary markedness constraints that require a final vowel in a surface form to be long if it is stressed (*/(mìni)si(tá)/).

What makes Tongan, as we analyze it, different from other cases in which loans import novel contrasts is that no novel structures in pronunciations result. Unlike the cases of 'special contrast' analyzed by Smith (2006), there is no need for otherwise undominated markedness constraints to be dominated in loans. Once words have been adapted, final vowel length is unproblematic for Tongan phonology: underlying forms for the words in question simply have a long or short final vowel, just like native words. Vowel deletion requires a novel contrast between full and weak vowels, but this contrast is realised only as a difference in the rates at which words vary between pronunciations that were already legal in native words (full vowel vs. devoiced or deleted vowel). The same underlying contrast between full and weak vowels is also realised as a difference in secondary stress position for certain words. If peninitial secondary stress were truly illegal in native Tongan words, then surface forms like /fa$_{weak}$(làki$_{weak}$)(séni$_{weak}$)/ 'fraction' would indeed be introducing novel surface structure. But as discussed in section 3.1, there seem to be no monomorphemic native words long enough to weigh in on this question; /fa$_{weak}$(làki$_{weak}$)(séni$_{weak}$)/ and /(tèmo)ka$_{weak}$(lási$_{weak}$)/ 'democracy' both have stress patterns that are legal in compounds, and whose legality in monomorphemic words is unknown, presumably even to learners.

From a study of just one language, we can't draw cross-linguistic conclusions. But it is suggestive that Tongan loans show sensitivity to non-native contrasts, despite how hard non-native contrasts are to perceive, and that this sensitivity does not result in previously-illegal surface structures. We speculate that non-native contrasts may be easier to encode in loans when this condition holds.

## 8 References


Alderete, John (1999). Head dependence in stress-epenthesis interaction. In Ben Hermans & Marc van Oostendorp (eds.) *The derivational residue in Optimality Theory*. Amsterdam: Benjamins. 29-50.

Anderson, Victoria & Yuko Otsuka (2006). The phonetics and phonology of "definitive accent" in Tongan. *Oceanic Linguistics* 45. 21–42.

Árnason, Kristján (1996). How to meet the European standard: word stress in Faroese and Icelandic. *Nordlyd: Tromsø University Working Papers on Language and Linguistics* 24. 1–22.

Bates, Douglas, Martin Maechler, Ben Bolker & Steven Walker (2014). lme4: Linear mixed-effects models using Eigen and S4. (R package version 1.1-6). http://CRAN.R-project.org/package=lme4.

Benua, Laura (1997). Transderivational identity: phonological relations between words. PhD dissertation, University of Massachusetts, Amherst.

Bermudez-Otero, Ricardo (2012). The architecture of grammar and the division of labour in exponence. In Jochen Trommer (ed.) *The morphology and phonology of exponence (Oxford Studies in Theoretical Linguistics 41)*. Oxford: Oxford University Press. 8-83.

Boersma, Paul (1998). *Functional phonology: formalizing the interaction between articulatory and perceptual drives*. The Hague: Holland Academic Graphics.

Boersma, Paul (2011). A programme for bidirectional phonology and phonetics and their acquisition and evolution. In Anton Benz & Jason Mattausch (eds.) *Bidirectional Optimality Theory.* Amsterdam: John Benjamins. 33-72.

Boersma, Paul & Silke Hamann (2009). Loanword adaptation as first-language phonological perception. In Andrea Calabrese & Leo Wetzels (eds.) *Loanword phonology*. Amsterdam: John Benjamins. 11-58.

Boersma, Paul & David Weenink (2006). Praat: Doing phonetics by computer (version 4.4). http://www.praat.org/.

Broselow, Ellen (1982). On predicting the interaction of stress and epenthesis. *Glossa* 16. 115–132.

Broselow, Ellen (2008). Stress-epenthesis interactions. In Bert Vaux & Andrew Nevins (eds.) *Rules, constraints, and phonological phenomena*. Oxford: Oxford University Press. 121-148.

Broselow, Ellen (2009). Stress adaptation in loanword phonology: perception and learnability. In Paul Boersma & Silke Hamann (eds.) *Phonology in perception*. Berlin: Mouton de Gruyter. 191-234.

Broselow, Ellen (2015). The typology of position-quality interactions in loanword vowel insertion. In Yuchau E. Hsiao & Lian-Hee Wee (eds.) *Capturing phonological shades within and across languages*. Newcastle Upon Tyne: Cambridge Scholars Publishing. 292-319.

Burzio, Luigi (1999). Surface-to-surface morphology: when your representations turn into constraints. Ms, Johns Hopkins University.

Cabré, Teresa (2009). Vowel reduction and vowel harmony in Eastern Catalan loanword phonology. In Marina Vigário, Sónia Frota & M. João Freitas (eds.) *Phonetics and phonology: interactions and interrelations*. Amsterdam and Philadelphia: John Benjamins. 267-286.

Churchward, C. Maxwell (1953). *Tongan grammar*. London: Oxford University Press.

Churchward, C. Maxwell (1959). *Tongan dictionary*. London: Oxford University Press.

Crosswhite, Katherine (1998). Segmental vs. prosodic correspondence in Chamorro. *Phonology* 15. 281–316.

Daland, Robert, Mira Oh & Syejeong Kim (2015). When in doubt, read the instructions: orthographic effects in loanword adaptation. *Lingua* 159. 70–92.

Davidson, Lisa (2006). Phonology, phonetics, or frequency: influences on the production of non-native sequences. *JPh* 34. 104–137.

Davidson, Lisa (2007). The relationship between the perception of non-native phonotactics and loanword adaptation. *Phonology* 24. 261–286.

Davidson, Lisa (2010). Phonetic bases of similarities in cross-language production: evidence from English and Catalan. *JPh* 38. 272–288.

Davidson, Lisa & Jason A. Shaw (2012). Sources of illusion in consonant cluster perception. *JPh* 40. 234–248.

de Lacy, Paul (2007). The interaction of tone, sonority, and prosodic structure. In Paul de Lacy (ed.) *The Cambridge handbook of phonology*. Cambridge University Press. 281-307.

Dupoux, Emmanuel, Takao Fushimi, Kazuhiko Kakehi & Jacques Mehler (1999). Prelexical locus of an illusory vowel effect in Japanese. In Krisztina Erdőhegyi, Géza Németh & Gábor Olaszy (eds.) *Proceedings of Sixth European Conference on Speech Communication and Technology, EUROSPEECH 1999*. Budapest: ISCA. 1675-1678.

Dupoux, Emmanuel, Kazuhiko Kakehi, Yuki Hirose, Christophe Pallier & Jacques Mehler (1999). Epenthetic vowels in Japanese: a perceptual illusion? *Journal of Experimental Psychology: Human perception and performance* 25. 1568–1578.

Dupoux, Emmanuel, C. Pallier, Núria Sebastián-Gallés & J. Mehler (1997). A destressing "deafness" in French? *Journal of Memory and Language* 36. 406–421.

Dupoux, Emmanuel, Sharon Peperkamp & Núria Sebastián-Gallés (2001). A robust method to study stress "deafness." *JASA* 110. 1606–1618.

Dupoux, Emmanuel, Sharon Peperkamp & Núria Sebastián-Gallés (2010). Limits on bilingualism revisited: Stress 'deafness' in simultaneous French–Spanish bilinguals. *Cognition* 114. 266–275.

Dupoux, Emmanuel, Núria Sebastián-Gallés, Eduardo Navarrete & Sharon Peperkamp (2008). Persistent stress "deafness": the case of French learners of Spanish. *Cognition* 106. 682–706.

Feldman, Harry (1978). Some notes on Tongan phonology. *Oceanic Linguistics* 17. 133–139.

Féry, Caroline (1991). German schwa in prosodic morphology. *Zeitschrift für Sprachwissenschaft* 10. 65–85.
Garellek, Marc & James White (2010). The acoustic correlates of stress and their use in diagnosing syllable fusion in Tongan. *UCLA Working Papers in Phonetics* 108. 35–65.
Garellek, Marc & James White (2015). Phonetics of Tongan stress. *Journal of the International Phonetic Association* 45. 13–34.
Gelman, Andrew, Yu-Sung Su, Masanao Yajima & Jennifer Hill (2010). arm: data analysis using regression and multilevel/hierarchical models. http://CRAN.R-project.org/package=arm.
Goldwater, Sharon & Mark Johnson (2003). Learning OT constraint rankings using a Maximum Entropy model. In Jennifer Spenader, Anders Eriksson & Östen Dahl (eds.), *Proceedings of the Stockholm Workshop on Variation within Optimality Theory*. Stockholm: Stockholm University. 111-120.
Hall, Kathleen Currie (2013). A typology of intermediate phonological relationships. *The Linguistic Review* 30. 215–275.
Hall, Nancy (2003). Gestures and segments: vowel intrusion as overlap. PhD dissertation, University of Massachusetts, Amherst.
Hall, Nancy (2006). Cross-linguistic patterns of vowel intrusion. *Phonology* 23. 387–429.
Hayes, Bruce (1995). *Metrical Stress Theory: principles and case studies*. Chicago: The University of Chicago Press.
Hayes, Bruce & Colin Wilson (2008). A Maximum Entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39. 379–440.
Hayes, Bruce, Colin Wilson & Ben George (2009). Maxent grammar tool. http://www.linguistics.ucla.edu/people/hayes/MaxentGrammarTool/.
Hillenbrand, James, Laura A. Getty, Michael J. Clark & Kimberlee Wheeler (1995). Acoustic characteristics of American English vowels. *JASA* 97. 3099–3111.
Hothorn, Torsten, Frank Bretz & Peter Westfall (2008). Simultaneous inference in general parametric models. *Biometrical Journal* 50. 346–363.
Hox, Joop J., Mirjam Moerbeek & Rens van de Schoot (2010). *Multilevel Analysis: Techniques and Applications*, Second Edition. New York: Routledge.
Inkelas, Sharon, C. Orhan Orgun & Cheryl Zoll (1996). Exceptions and static phonological patterns: cophonologies vs. prespecification. Ms, UC Berkeley and University of Iowa.
Iverson, Paul & Patricia K. Kuhl (1995). Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling. *JASA* 97. 553–562.
Iverson, Paul & Patricia K. Kuhl (1996). Influences of phonetic identification and category goodness on American listeners' perception of /r/ and /l/. *JASA* 99. 1130–1140.
Jäger, Gerhard (2007). Maximum Entropy Models and Stochastic Optimality Theory. In Jane Grimshaw, Joan Maling, Chris Manning, Jane Simpson & Annie Zaenen (eds.) *Architectures, rules, and preferences: a Festschrift for Joan Bresnan*. Stanford, CA: CSLI. 467-479.
Johnson, Keith (1997). Speech perception without speaker normalization: an exemplar model. In Keith Johnson & John W Mullenix (eds.) *Talker variability in speech processing*. San Diego: Academic Press. 145-165.
Kawahara, Shigeto (2011). Experimental approaches in theoretical phonology. In Marc van Oostendorp, Colin Ewen, Elizabeth Hume & Keren Rice (eds.) *The Blackwell companion to phonology*. Wiley-Blackwell. 2283-2303.

Kenstowicz, Michael (1996). Base-identity and uniform exponence: alternatives to cyclicity. In Jacques Durand & Bernard Laks (eds.) *Current trends in phonology: models and methods*. Manchester, England: European Studies Research Institute, University of Salford. 363-393.

Kenstowicz, Michael (1997). Quality-sensitive stress. *Rivista di Linguistica* 9. 157–187.

Kenstowicz, Michael. 2005. The phonetics and phonology of Korean loanword adaptation. In Sang-Jik Rhee (ed.) *Proceedings of the first European conference on Korean linguistics*. Seoul: Hankook. 17-32.

Kenstowicz, Michael (2007). Salience and similarity in loanword adaptation: a case study from Fijian. *Language Sciences* 29. 316–340.

Kinkade, M. Dale (1998). How much does a schwa weigh? In Ewa Czaykowska-Higgins & M. Dale Kinkade (eds.) *Salish languages and linguistics: theoretical and descriptive perspectives*. Berlin and New York: Mouton de Gruyter. 197-218.

Kubozono, Haruo (2006). Where does loanword prosody come from? A case study of Japanese loanword accent. *Lingua* 116. 1140–1170.

LaCharité, Darlène & Carole Paradis (2002). Addressing and disconfirming some predictions of phonetic approximation for loanword adaptation. *Langues et linguistique* 28. 71–91.

Lal, Brij V. & Kate Fortune (eds.) (2000). *The Pacific islands: an encyclopedia*. Honolulu: University of Hawaii Press.

Lewis, M. Paul (ed.) (2009). *Ethnologue: languages of the world*. 16th ed. Dallas, TX: SIL International.

Martin, Andrew (2005). Loanwords as pseudo-compounds in Malagasy. In Jeffrey Heinz & Dimitris Ntelitheos (eds.) *Proceedings of the 12th annual conference of the Austronesian Formal Linguistics Association*. Los Angeles: UCLA Working Papers in Linguistics. 287-295.

Mascaró, Joan (2002). El sistema vocàlic. Reducció vocàlica. In Joan Solà, Maria-Rosa Lloret, Joan Mascaró & Manuel Pérez Saldany (eds.) *Gramàtica del català contemporani*, vol. I. Barcelona: Empúries. 89-123.

McCallum, Andrew, Dayne Freitag & Fernando Pereira (2000). Maximum Entropy Markov Models for information extraction and segmentation. *Proceedings of the Seventeenth International Conference on Machine Learning*. San Francisco: Morgan Kaufmann. 591-598.

McCarthy, John J. & Alan Prince (1993). Generalized Alignment. In Geert Booij & Jaap van Marle (eds.) *Yearbook of Morphology*. Dordrecht: Kluwer. 79–154.

McCarthy, John J. & Alan Prince (1995). Faithfulness and reduplicative identity. In Laura Walsh Dickey, Jill Beckman & Suzanne Urbanczyk (eds.) *University of Massachusetts Occasional Papers in Linguistics*. Amherst, MA: GLSA Press. 249-384.

Mester, Armin (1992). Some remarks on Tongan stress. Ms., University of California, Santa Cruz.

Morton, E. J. (1962). A descriptive grammar of Tongan (Polynesian). PhD dissertation, Indiana University.

Munro, Pamela & Jason Riggle (2004). Productivity and lexicalization in Pima compounds. *BLS* 30. 114-126.

Pater, Joe (2000). Non-uniformity in English secondary stress: the role of ranked and lexically specific constraints. *Phonology* 17. 237–274.

Pater, Joe (2006). The locus of exceptionality: morpheme-specific phonology as constraint indexation. In Leah Bateman, Michael O'Keefe, Ehren Reilly & Adam Werle (eds.) *Papers in Optimality Theory III*. Amherst, MA: GLSA. 259-296.

Pater, Joe (2009). Morpheme-specific phonology: constraint indexation and inconsistency resolution. In Steve Parker (ed.) *Phonological argumentation: essays on evidence and motivation (Advances in Optimality Theory)*. London: Equinox. 123-154.

Peperkamp, Sharon (2005). A psycholinguistic theory of loan adaptations. In Marc Ettlinger, Nicholas Fleischer & Mischa Park-Doob (eds.) *BLS* 30. Berkeley, CA: Berkeley Linguistics Society. 341-352.

Peperkamp, Sharon & Emmanuel Dupoux (2003). Reinterpreting loanword adaptation: the role of perception. In Maria-Josep Solé, Daniel Recasens & Joaquin Romero (eds.), *Proceedings of the 15th International Congress of Phonetic Sciences*. Adelaide: Causal Productions. 367-370.

Peterson, Tyler (2007). Minimality and syllabification in Kabardian. *CLS* 39. Chicago: University of Chicago. 215-235.

Pierrehumbert, Janet (2001). Exemplar dynamics: word frequency, lenition, and contrast. In Joan Bybee & Paul Hopper (eds.) *Frequency effects and the emergence of linguistic structure*. Amsterdam: John Benjamins. 137-157.

Poser, William (1985). Cliticization to NP and Lexical Phonology. *WCCFL* 4. 262–272.

Prince, Alan (1983). Relating to the grid. *LI* 14. 19–100.

Prince, Alan & Paul Smolensky (1993). *Optimality Theory*. Malden, MA & Oxford: Blackwell.

R Core Team (2014). R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. www.R-project.org.

Round, Erich (2017). Phonological exceptionality is localized to phonological elements: the argument from learnability and Yidiny word-final deletion. In Claire Bowern, Laurence Horn & Raffaella Zanuttini (eds.) *On looking into words (and beyond): structures, relations, analyses*. Berlin: Language Sciences Press. 59-98.

Sailor, Craig (2010). Notes on Tongan vowel devoicing and deletion. Ms, University of California, Los Angeles.

Schütz, Albert J. (1970). Phonological patterning of English loan words in Tongan. In Stephen Adolphe Wurm & Donald C. Laycock (eds.) *Studies in honour of Arthur Capell (Pacific Linguistics C–13)*. Canberra: Australia National University. 409-428.

Schütz, Albert J (1978). English loanwords in Fijian. In Albert J Schütz (ed.) *Fijian language studies: borrowing and pidginization (Bulletin of the Fiji Museum 4)*. Suva: Fiji Museum. 1-50.

Schütz, Albert J (2001). Tongan accent. *Oceanic Linguistics* 40. 307–323.

Smith, Jennifer (2006a). Loan phonology is not all perception: evidence from Japanese loan doublets. In Timothy J. Vance & Kimberly A. Jones (eds.) *Japanese/Korean Linguistics 14*. Stanford: CSLI. 63-74.

Smith, Jennifer (2006b). Correspondence theory vs. Cyclic OT: beyond morphological derivation. In Chris Davis, Amy Rose Deal & Youri Zabbal (eds.) *NELS 36*. Amherst, MA: GLSA. 531-545.

Snijders, Tom & Roel Bosker (2011). *Multilevel analysis: an introduction to basic and advanced multilevel modeling*, 2nd ed. Los Angeles: SAGE Publications Ltd.

Steriade, Donca (2000). Paradigm uniformity and the phonetics-phonology boundary. In Janet Pierrehumbert & Michael Broe (eds.) *Acquisition and the Lexicon (Papers in Laboratory Phonology 5)*. Cambridge: Cambridge University Press. 313-334.

Stodden, Victoria, Friedrich Leisch & Roger D Peng (eds.) (2014). *Implementing Reproducible Research*. Boca Raton, FL: Chapman and Hall/CRC.

Taumoefolau, Melenaite (1998). *Problems in Tongan lexicography*. PhD dissertation, University of Auckland.

Taumoefolau, Melenaite (2002). Stress in Tongan. In Andrea Rackowski & Norvin Richards (eds.) *Proceedings of AFLA 8: the Eighth Meeting of the Austronesian Formal Linguistics Association (MIT Working Papers in Linguistics 44)*. 341–354.

Uffmann, Christian (2007). *Vowel epenthesis in loanword adaptation*. Tubingen: Max Niemayer Verlag.

Välimaa-Blum, Riitta (1986. Finnish vowel harmony as a prescriptive and descriptive rule: an autosegmental account. In Fred Marshall, Ann Miller & Zheng-sheng Zhang (eds.) *Proceedings of the third Eastern States Conference on Linguistics*. Pittsburgh: University of Pittsburgh. 511-522.

Vicenik, Chad & Grace Kuo (2010). Tongan intonation. Poster presented at the 160th Meeting of the Acoustical Society of America, Cancún.

Zuraw, Kie (2015). Allomorphs of French *de* and their ability to take wide scope in coordination. *Linguistics Vanguard* 1.

Zuraw, Kie & Kevin Huynh (submitted). Solving the problem of whole-language simulation: a case study from Tongan loan adaptation.